AUSTRALIAN JOURNAL OF PHYSICS

PAPERS FROM
CONFERENCE ON SOLID STATE PHYSICS
MELBOURNE, AUGUST 1959

VOLUME 13, NUMBER 2A, JULY 1960

REGISTERED IN AUSTRALIA FOR TRANSMISSION BY POST AS A PERIODICAL

AUSTRALIAN JOURNAL OF PHYSICS

Published by the Commonwealth Scientific and Industrial Research Organization. Volumes 1 to 5 of the Australian Journal of Physics and the Australian Journal of Chemistry issued as the Australian Journal of Scientific Research, Series A: Physical Sciences. Issued quarterly, £2 per annum.

BOARD OF STANDARDS

The Board of Standards for this Journal is appointed by the Commonwealth Scientific and Industrial Research Organization and the Australian Academy of Science and consists of Professor Sir Macfarlane Burnet (Chairman), Dr. N. S. Noble (Editor), Professor N. S. Bayliss, Dr. J. L. Pawsey, Professor W. P Rogers, and Professor J. S. Turner.

ADVISORY COMMITTEE

Acceptance of papers for this Journal is in the hands of an Advisory Committee appointed by the Board of Standards in consultation with the Institute of Physics (Australian Branch) and consisting of Dr. N. S. Noble (Chairman and Editor), Dr. G. H. Briggs, Emeritus Professor L. G. H. Huxley, Dr. J. L. Pawsey, and Professor H. C. Webster.

OTHER JOURNALS PUBLISHED BY C.S.I.R.O.

Australian Journal of Chemistry Australian Journal of Biological Sciences Australian Journal of Applied Science	Issued quarterly, £2 per annum
Australian Journal of Agricultural Research	Issued six times a year, £2 per annum
Australian Journal of Marine and Freshwater Research Australian Journal of Botany Australian Journal of Zoology	Not issued at regular intervals, 10/- per issue

The Organization is a signatory to the Fair Copying Declaration, details of which may be obtained from the office of the Royal Society, London.

All enquiries and manuscripts should be forwarded to:

The Editor,
Australian Journal of Physics,
Commonwealth Scientific and Industrial Research Organization,
314 Albert Street, East Melbourne, C.2, Victoria

MELBOURNE

CONFERENCE ON SOLID STATE PHYSICS

A conference on Solid State Physics was held in Melbourne from 17th to 21st August, 1959, under the auspices of the Australian Branch of the Institute of Physics. This conference was the first of its kind in Australia and attracted an attendance of about 130.

The Institute was assisted financially in holding this conference by grants from the Aeronautical Research Laboratories, Amalgamated Wireless (Australasia) Ltd., the Australian Atomic Energy Commission, the British General Electric Company Pty. Ltd., the Commonwealth Scientific and Industrial Research Organization, Philips Electrical Industries Pty. Ltd., the Rola Company (Australia) Pty. Ltd., Standard Telephones and Cables Pty. Ltd., the Australian Academy of Science, the University of Melbourne, and the University of New South Wales.

Forty-six papers were read and discussed at the conference and a number of the original contributions are being published in this issue. No attempt was made to record the discussion at the conference but many authors have incorporated in their papers the results of such discussion. Brief summaries and notes on all papers presented have been published in the British Journal of Applied Physics (**11** : 137 (1960)).

The conference was organized on behalf of the Institute of Physics by the following Committee: W. Boas, R. I. Garrod, P. G. Klemens, J. F. Nicholas, A. L. G. Rees, and S. E. Williams.

THE STRUCTURE OF BISMUTH AT LOW TEMPERATURES

By C. S. Barrett*

[*Manuscript received September* 25, 1959]

Summary

A precision determination of the displacement parameter of zone-refined bismuth containing less than 0·3 parts per million of spectroscopically detectable impurity has been carried out as a function of temperature. Procedures were used that minimize uncertainties in integrated intensities due to extinction, absorption, temperature factor, dispersion and f-curve errors, and statistical errors. The parameter b/d, where d is the spacing of the (111) rhombohedral planes of one sub-lattice and b is the displacement of the other sub-lattice (111) planes from these, is $0 \cdot 40437 \pm 0 \cdot 00008$ at 4·2 °K, $0 \cdot 40403 \pm 0 \cdot 00008$ at 78 °K, and $0 \cdot 4041 \pm 0 \cdot 0005$ at 296 ± 2 °K; the rhombohedral parameter, u, corresponding to these values is, respectively, 0·23406, 0·23400, and 0·23401. The small but presumably significant dependence on temperature at low temperatures is in disagreement with predictions of a theory of Blount and Cohen and may signify a lessening of resistance to the atom displacement from simple cubic, due to thermal expansion of the structure. Powder patterns yielded $57 \cdot 31_4$, $57 \cdot 28_6$, and $57 \cdot 23_6°$ for the rhombohedral angle at 4·2, 78, and 304 °K respectively.

I. Introduction

There is increased interest in recent years in the semi-metal bismuth and the related elements of Group V_B, which have similar rhombohedral structures, and in the solid solutions of these elements. The crystal structures of these materials are not complex, but complicated effects on the electrical and magnetic properties are observed upon alloying. A number of investigations have been undertaken in this institute to reveal these effects in bismuth and its alloys and to relate the effects to the crystal structures and the energy band structures of these materials.

In discussing models that may explain these effects, it is useful first to take a highly simplified view of the bismuth structure. In a crude sense, the structure of bismuth can be regarded as nearly simple cubic. Is the fact that it is nearly cubic significant? Hume-Rothery and Raynor (1954) apparently feel that it is not. They point out that hybridization of the *sp* orbitals causes a departure of the bonds from the right-angled *p*-type positions they have in cubic crystals. Both the amount of *sp* hybridization and the observed distortion of the crystal structure from simple cubic does increase, to be sure, as one passes from bismuth to antimony to arsenic. However, *sp* hybridization in itself fails to account for the marked deviations from Vegard's law of the lattice parameters of the bismuth-antimony alloys, recently observed by Jain (1959), which suggest strongly that energy band considerations are important.

Blount and Cohen (personal communication 1958), in their theory of the band structure and crystal structure of bismuth, assign great significance to the

---

* Institute for the Study of Metals, University of Chicago, Chicago 37, Illinois, U.S.A.

small deviation of the structure from cubicity. They call attention to the fact that bismuth in its chemical behaviour is normally trivalent, despite the fact that there are two *s* and three *p* electrons outside of the closed shell. The explanation for trivalence instead of pentavalence is that the *sp* excitation energy in heavy atoms is very large so that the two *s* electrons act as an " inert pair ", as in thallium and lead. The three *p* electrons would be expected to provide bonds to nearest neighbours that stand at right angles to each other, forming a cubic structure. But if one now considers the *p* energy bands in the first Brillouin zone of the simple cubic structure, one sees that these bands can hold six electrons per atom (the factor of 2 coming from spin). Thus with only three electrons per atom, the zone in bismuth is only half filled.

Imagine now a small change in the crystal structure that increases the unit cell size by a factor of 2 and decreases the Brillouin zone size by the same factor. This new Brillouin zone could contain exactly three electrons per atom. Moreover, the effect of the change in structure would be to lower the energy of the occupied portion of the band and to raise the energy of the unoccupied portion. Such a distortion could in principle stabilize the new crystal structure.

The actual structure of bismuth is just such a one as could give stabilization according to this model of Blount and Cohen. The structure consists of two interpenetrating face-centred sub-lattices, one of which is displaced from mid point positions of the other by a small amount. It may be likened to the structure of sodium chloride with the chlorine sub-lattice displaced from its normal position within the sodium sub-lattice by a few per cent. along the axis of 3-fold symmetry (and with the interaxial angle reduced a few degrees). The Blount-Cohen model has been elaborated to give a similar understanding of the crystal structure of polonium, tellurium, and selenium, and to explain observed features of the band structure of bismuth.

Let us now examine the predictions of this model regarding the crystal structure of alloys and the effect of temperature on the structure of bismuth. Consider first the alloying of bismuth with an acceptor such as lead. As lead is added, electrons will primarily be removed from the valence band rather than from the conductance band because of the high density of states in the valence band. Since the states in the valence band are lowered in energy by the displacement of the sub-lattices, it follows that removing electrons from these states should decrease the distortion. The state of development of the theory is not adequate to give even an estimate of the magnitude of the effect, so one must be prepared to measure even small changes.

Turning next to a contrasting situation, if a donor such as tellurium is progressively added to bismuth the lowered states are increasingly occupied, and hence the displacement should increase—at least initially. When all the holes in the valence band are filled, however, the added electrons will begin to enter the states in the conduction band, which are states that have been raised in energy by the displacement, and the trend of the displacement should reverse and decrease toward cubicity. Work leading to a test of these predictions is currently under way.

The present experiments concern the effect of temperature on the structure of extremely pure bismuth. On the basis of the model discussed above, thermal excitation of electrons from the lowered states of the valence band to the raised states of the conduction band should act to decrease the displacement of the sub-lattices. The changes expected upon raising the temperature should be smaller, in general, than would be expected upon alloying, because the number of electrons affected is very much smaller. Near room temperature the number of electrons that would be excited is of the order of the number that would be removed by adding 0·01 per cent. of lead (Blount and Cohen, personal communication 1958).

It will be noted that stabilization of the bismuth structure is associated in this model with the displacement parameter rather than with the distortion of the interaxial angle from that corresponding to simple cubic, which is viewed by Blount and Cohen as a secondary effect. Accordingly, the attention in the present work is primarily on the displacement parameter, though some data on the rhombohedral angle distortion are included. Effort is directed toward the attainment of precision in the parameter determination, both for the purpose of disclosing any thermal effects and for later comparison with bismuth alloy data; operating at low temperatures is, of course, a prerequisite for maximum precision, and the use of extremely pure crystals is demanded in order to avoid alloying effects. The best of the previous data, that of Goetz and Hergenrother (1932), is wholly inadequate for present purposes.

II. MATERIAL

A sample of extremely high purity bismuth that had been zone-refined for ultrasonic attenuation measurements was kindly furnished by D. H. Reneker. It has been prepared from Cerro de Pasco Corporation bismuth of 99·998 per cent. purity, by zone refining in a vacuum, using a boat that had been machined of "AUC" graphite and subsequently boiled in aqua regia. Ingots about $\frac{3}{4}$ by $\frac{3}{4}$ by 20 in. were given at least 15 passes, with additional passes being made to grow crystals of suitable size.

The sample was taken from near the mid length of the ingot. Spectroscopic analysis of this portion disclosed no detectable impurities except Ag and Cu, and the intensity of the lines from these elements was less than that of the lines from a standard sample that contained 3 parts in 10^7 of Cu and Ag. A gas analysis indicated that the concentration of dissolved gases was less than 1 part in 10^7 by weight. Electrical measurements at helium temperature indicated that the crystals had an electronic mean free path of at least 1 mm. The ratio of resistance of a single crystal specimen at 298 and 1·2 °K was 220.

Crystals about 1 cm^3 in volume were cut from the ingot by a string saw, using "Saran" braid dipping in a nitric acid solution, the sample being warmed to about 40 °C during the cutting. After orienting the crystals by back reflection Laue photographs, faces were cut normal to the 3-fold axis in a surface grinder using a water-cooled high speed alundum cut-off wheel that was advanced 0·0005 in. per pass. Subsequent electropolishing reduced the amount of distortion in the surface layer to the desired level. Aid in maintaining a flat surface during the electropolishing was obtained by alternating brief electropolishing and gentle, uniform abrasion by alundum powder in an air jet. On one sample a controlled degree of imperfection was purposely left on the surface by carrying the electropolishing just to the point of removal of substantially all of the small, uniformly sized twins that were introduced by the alundum. There were no resolved $K\alpha$ doublets in the diffraction pattern from this crystal. A second crystal, prepared with deeper electropolishing, gave resolved doublets.

III. METHODS

Integrated intensities were measured with a low temperature diffraction unit consisting of a cryostat mounted on a spectrometer and fitted with a flat-crystal monochromator, synchronous-motor drive for rotation of crystal and counter, and scintillation counter with pulse-height analyser to remove fluorescent and half-wavelength components of the radiation. The apparatus provided continuous recording in order to detect sudden changes in the functioning of the apparatus, but all counts were taken from a scaler connected in parallel with the recorder. Electronic voltage stabilizers were employed fcr all X-ray and counting circuits.

Background corrections were made in the usual way, averaging the intensities measured at points of minimum intensity on each side of a peak. Sudden changes in background, abnormally high background, or significant changes in measured intensities in check tests during a run were taken as grounds for discarding the run. Normal runs at low temperatures required 4–8 hr, but conclusions were drawn from pairs of peaks that were measured one directly after the other during a run, thus minimizing the effect of any slow drifts in the apparatus.

The integrated intensities of reflections depend not only upon the parameter that is to be determined but also upon factors involving extinction, absorption, atomic scattering factors, anomalous dispersion, and Debye characteristic temperature ; each of these factors requires attention if errors are to be minimized. As rigorous treatments of some of these are not available, procedures were designed so that the uncertainties introduced by them would be small, particularly with regard to the relative values of the parameter at different temperatures. Since the effects sought required the attainment of the highest precision, these factors are discussed in some detail.

It was desirable to use single-crystal rather than powder measurements, in order to gain intensity in the higher orders and to minimize overlapping reflections. This made the extinction correction a matter of major concern. The possibility of removing extinction by cold working the crystals was explored, but because plastic flow in bismuth is predominantly by deformation twinning, even with small strains, this approach proved unprofitable : suitably high densities of residual dislocations were not reached without a high percentage of the crystals being twinned.

Extinction was eliminated as a factor by basing the parameter determination on weak reflections. The 111, 222, and 333 reflections (rhombohedral indices) were much too strong to use, and there was danger that the somewhat weaker 555 may have shown some extinction effects ; the background intensity began to rise somewhat as the diffraction angle was reduced below that for the fifth order, suggesting that a greater uncertainty in the background correction would be encountered in this range. The 666, 777, 888, and 999 reflections were best for the purpose at low temperatures but 999 was too weak to use at room temperature. Higher orders were also obtained at low temperatures, the tenth and twelfth being of moderate intensity and the eleventh being seen as a trace on some runs, but the alignment of these in the receiving slits was uncertain so they were not used in the analysis. Since the integrated counts from the chosen reflections were rather small (usually 1000 to 10,000) it was important to average the results of a number of runs. Since a theoretical calculation of background intensities for this crystal structure is not available, a linear background was assumed under each peak.

The choice of reflections to be used was governed not only by the extinction, background, and alignment considerations just mentioned but also by a desirable balancing of errors (discussed below) that occurs when these adjacent high-order reflections are used. This choice of orders is also desirable to minimize and to balance out uncertainties in the absorption factor, which might be serious at low orders.

No absorption factor correction enters into the parameter determination if the beam strikes a crystal surface that is parallel to the reflecting plane and mirror-smooth. Care was taken in the cutting and electropolishing to meet this condition as closely as possible, and in aligning the crystal so as to keep the irradiated spot (approximately 1 by 2 mm) away from all edges of the prepared surface (approximately 8 by 10 mm). Runs were made with a crystal in one position and then with the same crystal after being rotated 180° about the normal to the surface.

Crystals were inspected by X-ray reflection micrographs (Schulz method) and only those surfaces were used that were free from misoriented subgrains and that had reasonably uniform reflecting power as imaged in the Schulz pictures. However, the pictures contained a few narrow, parallel dark lines that appeared to be caused by residual rows of dislocations left from the cutting wheel and not fully removed in electropolishing. To reduce the effect of any unavoidable variation in conditions over the surface of the crystals, runs were made with the beam striking different areas and the results were averaged. By basing results on pairs of peaks adjacent to each other, the sixth paired with the seventh, etc., the area irradiated was kept nearly the same for the two measurements of each pair in each run.

The incident beam was closely defined by slits at the monochromator, but the diffracted beam entered the counter through wide slits that were opened enough to accept the entire peak as the crystal turned through the reflecting range in the moving-crystal, moving-counter technique. Care was taken to limit the vertical length of the collimated beam so that no part of the incident or the reflected beam struck the edges of the beryllium windows and thermal shields of the cryostat or the slits of the counter.

To interpret the data the usual formula was used for the integrated intensity, $E\omega/I$:

$$E\omega/I = Ae^{-2M} f^2 L(1+\cos 2\pi nb/d),$$

where A is a constant, $\exp(-2M)$ is the Debye-Waller temperature factor, f is the scattering factor, n the order of reflection, b is the displacement of atoms of one sub-lattice with respect to the other in the direction of the 3-fold axis expressed in terms of the parameter b/d where d is the spacing of the (111) planes (rhombohedral axes); referred to hexagonal axes, $d=\frac{1}{3}c$. The parameter b/d may be converted to the parameter u of the smallest rhombohedral cell by the relationship $u=\frac{1}{6}(1+b/d)$. The Lorentz polarization factor L is $(1+\cos^2 2\theta_M \cos^2 2\theta)/\sin 2\theta$, where θ_M is the diffraction angle for the monochromator and θ that for the specimen; for Ag $K\alpha$ and the lithium fluoride monochromator used, $\cos^2 2\theta_M = 0\cdot92392$ and for Mo $K\alpha$ this term is $0\cdot87925$.

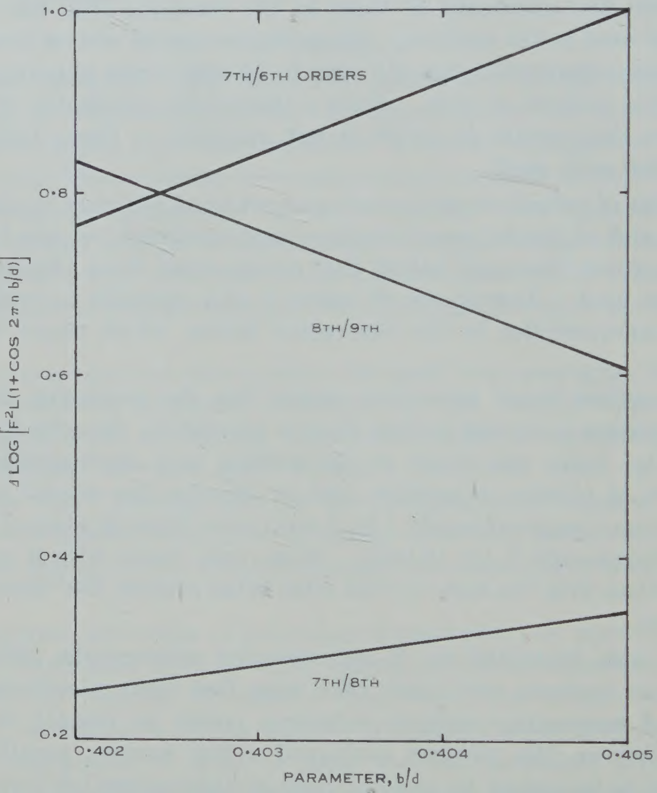


Fig. 1.—Curves for determining the parameter b/d of bismuth from integrated intensities of Ag $K\alpha$ reflections using ratios of 777/666, 888/999, and 777/888 (rhombohedral indices) after correction by appropriate factors.

To derive the parameter from pairs of reflections the logarithmic form was used:

$$\log(E\omega/I) + 2M \log e = \log A + \log\{f^2 L(1+\cos 2\pi nb/d)\}.$$

A curve was plotted for each pair of reflections, using the difference between the last term on the right for the two reflections as ordinate and b/d as abscissa, Figure 1. Then the difference between the magnitudes on the left-hand side of the equation for the same two reflections, when referred to the curve, yielded the parameter value.

A scattering factor curve for bismuth based on self-consistent field wave-functions is not available ; it was decided to base a curve on recent Hartree-Fock self-consistent field calculations with correction for anomalous dispersion due to the K electrons. The f-curve used constitutes a source of uncertainty, since the curve is not fully reliable, but it was found that the balancing of errors inherent in the method used in this work made this source of error a minor one. A more detailed discussion of this factor, the Debye-Waller temperature factor, and the diffraction angles used in the computations is included in Appendix I.

With regard to the Debye-Waller temperature factor, Goetz and Jacobs (1937) concluded from their measurements on bismuth over a range of temperatures that the temperature factor in the usual form with zero-point energy included suitably accounts for the temperature variation of intensities of reflections from the bismuth cleavage face, but the value deduced from their measurements, 95·9 °K, is markedly smaller than the values obtained by various investigators from specific heat measurements and was not used. Instead, the internal consistency of parameter determinations from different pairs of orders was used to determine an effective characteristic temperature (Appendix I) ; for diffraction at 78 °K this was found to be 127 °K and at room temperature it was found to be 124 °K. At 4·2 °K this method of determination would not be satisfactory because the temperature factor is so small. A value of 127 °K was arbitrarily assumed, but other reasonable values gave almost identical results.

IV. RESULTS

Eleven sets of readings consisting of measurements at 4·2 °K with Ag $K\alpha$ reflections in the sixth, seventh, eighth, and ninth orders, computed with the characteristic temperature $\Theta=127$ °K gave, on averaging and computing the standard deviation, $b/d=0\cdot40437\pm0\cdot00008$, a result quite insensitive to the temperature factor.

Ten sets of readings of the same orders measured at 78 °K and computed with $\Theta=127$ °K (derived from internal consistency in the data) gave $b/d=0\cdot40403\pm0\cdot00008$. Reasonable changes in Θ altered b/d less than the standard deviation. For example, using $\Theta=121$ °K gave $b/d=0\cdot40402$.

Seven sets of the sixth, seventh, and eighth orders measured at 294–298 °C and computed with the value of Θ indicated by internal consistency, 124 °K, gave $b/d=0\cdot40408\pm0\cdot00047$. Here the temperature factor is larger and the balancing of errors is less complete (an arbitrary change of Θ to 127 °K yielded $b/d=0\cdot40419\pm0\cdot00046$).

These values obtained with Ag $K\alpha$ radiation are plotted in Figure 2. It will be seen that b/d appears to be significantly lower at 78 °K than at 4·2 °K. If one assumes a normal distribution of errors in the measurements at these two temperatures and applies the usual formula for computing the difference of the mean values, it follows that

$$b/d_{4\cdot 2}-b/d_{78}=\{34\pm\sqrt{(8^2+8^2)}\}\times 10^{-5}=(34\pm 11)\times 10^{-5}.$$

Thus, with a difference three times the standard deviation, the difference, although small, more than meets the customary test for significance. On the other hand, the parameter at room temperature, with its greater scatter, does not differ significantly from the parameters at low temperatures. Averages for the more perfect crystal and for the less perfect crystal did not differ by a statistically significant amount, so the results from the two were combined in obtaining the values reported.

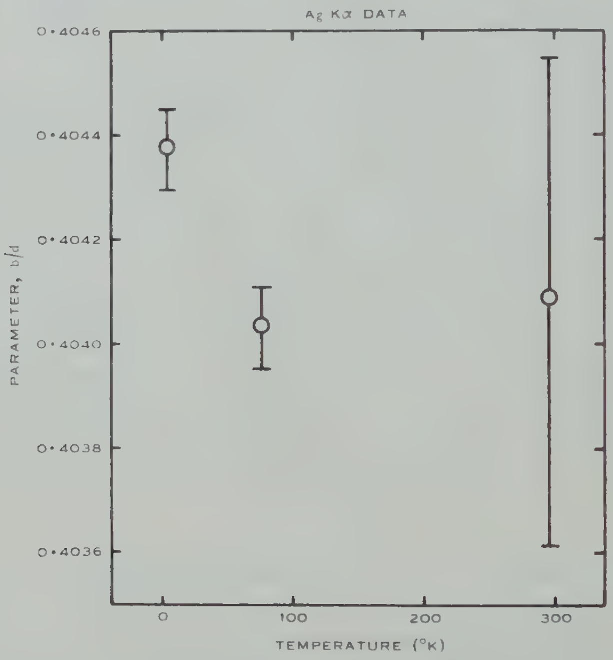


Fig. 2.—Values of the bismuth parameter b/d determined from averages of runs with Ag $K\alpha$ radiation. Error limits shown are computed standard deviation values of the individual determinations.

For comparison with results using Ag $K\alpha$, some preliminary runs were made with Mo $K\alpha$. Background intensities were higher and instrumental conditions were less stable, so the results were not regarded as comparable in reliability to those with Ag $K\alpha$—in fact were considered to be worthless at room temperature—but they serve as a rough check at low temperatures. Five sets of readings at 4·2 °K including the sixth, seventh, eighth, and ninth orders were computed, using $\Theta=107$ °K; the average gave $b/d=0\cdot 40412\pm 0\cdot 00020$. Four sets of readings of the same reflections taken at 78 °K and computed with the preferred value of $\Theta=127$ °K gave $b/d=0\cdot 40432\pm 0\cdot 00012$ (an arbitrary change to $\Theta=121$ °K yielded $b/d=0\cdot 40431\pm 0\cdot 00012$).

It is not clear what weight should be given these preliminary runs; if they are averaged together with the runs using Ag radiation, all runs being given equal weight, the grand averages become $0 \cdot 40429 \pm 0 \cdot 00007$ at 4·2 °K and $0 \cdot 40411 \pm 0 \cdot 00006$ at 78 °K. The difference between these is $(18 \pm 9) \times 10^{-5}$, a difference which again meets the usual test for significance at the 5 per cent. level and indicates the same direction of change as in Figure 2. It is our opinion that the results based on Ag radiation alone (Fig. 2) are probably more reliable than the equally weighted combined data.

The present results differ from the 1932 determination of the parameter by Goetz and Hergenrother, which yielded the room temperature value of $0 \cdot 409 \pm 0 \cdot 0007$. The latter was based mainly on the intensities of 333, 444, and 555, for 666 was not detected and 777 was given little weight; there is a possibility, therefore, that extinction may have affected these results. The results of Figure 2, $b/d=0 \cdot 40437$, 0·40403, and 0·40408 for 4·2, 78, and 294–298 °K respectively, are equivalent, respectively, to $u=0 \cdot 23406$, 0·23400, and 0·23401, and the Goetz-Hergenrother value is equivalent to 0·2348.

The results in Figure 2 can be restated as follows. The displacement of rhombohedral (111) planes of one of the sub-lattices from the mid point positions that correspond to cubicity are given as a fraction of the (111) periodicity by $\frac{1}{2}-b/d$; thus for 4·2, 78, and 298 °K the parameters quoted correspond to displacements of $0 \cdot 0956_3$, $0 \cdot 0959_7$, and $0 \cdot 0959_1$ respectively. Thus there is an increase of 0·35 per cent. in displacement in going from 4·2 to 78 °K.

It is interesting to compare this displacement with the deviation from cubicity in interaxial angles. Powder diffraction patterns were recorded, using the same material and the same diffractometer, but with narrower slits and with Cu $K\alpha$ filtered radiation. The attainment of extreme precision was not attempted in this part of the work.

Analysis of these patterns, using a Nelson-Riley plot in the range of reflections between $\Theta=46$ to 75°, gave c/a values (hexagonal axes) and estimated errors as follows:

at 4·2 °K, $c/a=2 \cdot 604_4 \pm 0 \cdot 0002$,
at 78 °K, $c/a=2 \cdot 606_2 \pm 0 \cdot 0003$,
at 304 °K (from Castelliz' work, see Appendix I)
$c/a=2 \cdot 609_1$.

The corresponding rhombohedral angles at 4·2, 78, and 304 °K were $57 \cdot 31_4$, $57 \cdot 28_6$, and $57 \cdot 23_6$° respectively. In a cubic crystal this angle would be 60°, thus the distortion from cubic was $2 \cdot 68_6$, $2 \cdot 71_3$, and $2 \cdot 76_4$° respectively. The hexagonal cell dimension a from these powder patterns was $4 \cdot 533_3$ Å at 4·2 °K and $4 \cdot 534_5$ Å at 78 °K with estimated errors of about $\pm 0 \cdot 0005$ Å.

The powder patterns, which were prepared after simple cooling of annealed powders in the cryostat, showed no evidence of a transformation of phase.

V. DISCUSSION

As expected from the small number of electrons thermally excited, the effect of temperature on the displacement parameter is exceedingly small. The model of Blount and Cohen mentioned in the introduction implies a *decrease* in the displacement from cubicity, whereas the experimental results indicate a slight *increase* in raising the temperature from 4·2 to 78 °K, so a modification of the model is called for.

One might imagine that there is a relatively low potential barrier separating an atom that is resting on *one* side of the mid point between neighbouring rhombohedral (111) planes from a possible position on the *other* side, and that because of this the excursions made by the atom during thermal vibrations would be greater toward the mid point than toward the nearest (111) plane. But this factor would act to *decrease* the displacement as the temperature is raised, and thus, like the band model discussed above, if this factor has any appreciable magnitude it is overcome by some other factor.

A possible factor acting in the right direction is proposed by Cohen (personal communication, 1959). Because of the small number of electrons that are thermally excited, it can be assumed that the driving force for the displacement is effectively independent of temperature in the temperature range investigated. But, on the other hand, perceptible thermal expansion occurs. This would be expected to lessen the resistance of the crystal structure to the displacement as the temperature is raised, and with constant driving force should give a tendency toward *increased* displacement at the higher temperatures.

VI. Acknowledgments

It is a pleasure to thank Professor M. H. Cohen (Institute for the Study of Metals, University of Chicago) for proposing the work, Professor C. B. Walker (Institute for the Study of Metals) for helpful discussions, Dr. D. Reneker (former graduate student, University of Chicago) for purified bismuth, Dr. Lotte K. H. Castelliz and Dr. H. Christine Sosnovsky (chemists, formerly at the Institute for the Study of Metals) for aid in computation, and the staff of the cryogenic laboratory (University of Chicago) for liquid helium. The work was supported in part by the Office of Naval Research, Contract No. Nonr 2121(11).

VII. References

Dauben, C. H., and Templeton, D. (1955).—*Acta Cryst.* **8** : 841.

Freeman, A. J. (1959).—*Acta Cryst.* **12** : 261.

Goetz, A., and Hergenrother, R. C. (1932).—*Phys. Rev.* **40** : 137.

Goetz, A., and Jacobs, R. B. (1937).—*Phys. Rev.* **51** : 151.

Hume-Rothery, W., and Raynor, G. V. (1954).—" The Structure of Metals and Alloys." 3rd Ed., p. 49. (Institute of Metals : London.)

International Tabellen zur Bestimmung von Kristallstrukturen (1935).—Vol. II. (Borntrager : Berlin.)

Jain, A. L. (1959).—*Phys. Rev.* **114** : 1518.

James, R. W. (1948).—" The Optical Principles of the Diffraction of X-rays." (Bell and Sons : London.)

James, R. W., and Brindley, G. W. (1931).—*Phil. Mag.* (7) **12** : 81.

Kalinkina, I. N., and Strelkov, P. G. (1958).—*Soviet Phys. JETP* **34** : 426.

Parratt, L. G., and Hempstead, C. F. (1954).—*Phys. Rev.* **94** : 1593.

Pearson, W. B. (1958).—" Handbook of Lattice Spacings and Structures of Metals and Alloys." (Pergamon Press : New York.)

Zener, C., and Bilinsky, S. (1936).—*Phys. Rev.* **50** : 101.

Appendix I

Scattering Factor Curve.—The *f*-curve used was based on the recent Hartree-Fock self-consistent field calculations of Freeman (1959) for Tl^+, introducing the change of scales of ordinates and abscissae to Bi (a change from 80 to 83 electrons) by the method of James and Brindley (1931). The orders used in the present work fell in a rather high range of $(\sin\theta)/\lambda$ where the *f*-curve is chiefly dependent on electron densities relatively deep in the atoms and where the curve changes shape but little with atomic number. (The $(\sin\theta)/\lambda$ values for the 666 to 999 reflections with Ag $K\alpha = 0 \cdot 560833$ Å and bismuth at 4·2 °K fall in the range 0·7629 to 1·1443). A Hönl correction for anomalous dispersion was applied to the *f*-curve. By choosing Ag $K\alpha$ radiation, the ratio of the incident wavelength to the wavelength of the K absorption edge was large (4·09) so that the K electron dispersion correction was very small; the correction applied was $-1 \cdot 1$, independent of diffraction angle. This is the real part of the K correction term, and, since the incident wavelength is larger than the K absorption limit, the imaginary part of the correction is zero (James 1948). The contribution of L and M electrons to the correction was ignored; since the ratio of the emission wavelength to the wavelength of the Bi L_{III} absorption edge is only 0·6, these contributions are also small (Parratt and Hempstead 1954). Judging by the comparison between theory and experiment with tungsten, the total correction at a wavelength of 0·6 of the L_{III} absorption edge can be ignored about as safely as it can be applied. In any event, uncertainties in the dispersion correction should be largely ironed out in averaging the pairs of reflections that were chosen. The consistency of the results for the different pairs indicates that the structure factors used were satisfactory. In preliminary runs with Mo $K\alpha$ radiation, the dispersion correction tabulated by Dauben and Templeton (1955) was used.

To test the effect of using a different *f*-curve on the Ag $K\alpha$ results, the older values based on James and Brindley's work (International Tabellen 1935) were plotted with a dispersion correction of 1·1, independent of angle, subtracted. The effect on the parameter b/d computed from a single pair of reflections using this curve was not negligible, but the balancing of errors when the three pairs of reflections were averaged indicated that this source of error was a minor one, altering the average b/d by only about 0·00002, as will be seen from Table 1.

Diffraction Angles.—Calculations were based on interplanar spacings, d_{111}, obtained by extrapolation of single-crystal data on a Nelson-Riley plot, with the Ag $K\alpha$ wavelength being taken as 0·560833 Å. The value $d_{111} = 3 \cdot 9324$ Å was used for 4·2 °K, and, although subsequent powder data indicated that 3·935 might be a better value, the results would not be significantly altered by using the latter. For 78 °K the spacing used was 3·9381 Å, which was the average of 3·9391 and 3·9372 respectively for the two crystals, and was satisfactorily checked by the powder data. For room temperature work 3·95407 Å was used (Pearson 1958).

The room temperature value was also checked in this laboratory by back reflection symmetrical focusing camera powder patterns analysed by L. Castelliz, using Cohen's least squares method and the same stock of zone-refined bismuth that was used in the single-crystal work. With hexagonal axes, the unit cell at 31 °C had dimensions $a=4\cdot54613\pm0\cdot00018$ Å, $c=11\cdot86152\pm0\cdot00060$ Å, $c/a=2\cdot6091$; and with rhombohedral axes, $a=4\cdot74568\pm0\cdot00020$ Å, $\alpha=57° 14' 11''\pm13''$. Here again the difference between these and the earlier values is not important in the determination of b/d.

The Effective Characteristic Temperature.—A rigorous temperature factor correction for bismuth is not available. It was decided to employ the usual formula that is intended to apply to cubic crystals (James 1948), which includes a zero-point energy term, and to use an *effective* Debye characteristic temperature in the formula.

TABLE 1

EFFECT ON b/d OF USING NEW AND OLD f-CURVE DATA WITH $\Theta=127$ °K

Temperature	Reflection Pair	Average Log Ratio of Intensity	b/d with Old Curve	b/d with New Curve
4·2 °K ..	777/666	0·91604	0·40422	0·40431
	777/888	0·33681	0·40424	0·40409
	888/999	0·65009	0·40460	0·40472
Average ..			0·40435	0·40437
78 °K ..	777/666	0·87327	0·40405	0·40409
	777/888	0·36888	0·40410	0·40393
	888/999	0·74212	0·40389	0·40407
Average ..			0·40401	0·40403

At first an attempt was made to use characteristic temperature values near those obtained from the specific heat work. Modification of specific heat values for this purpose is discussed by Zener and Bilinsky (1936) for isotropic crystals, but a detailed treatment for bismuth is not available; as a first approximation, characteristic temperatures Θ of 107 and 121 °K were chosen for use with the data taken at 4·2 and 78 °K respectively, and 121 °K was also chosen for the room temperature data, these values being about 3 per cent. above some recent specific heat values but scarcely out of the scatter band of reported values (Kalinkina and Strelkov 1958). (In the simplified cases usually discussed, the X-ray value is expected to be a few per cent. higher than is obtained from specific heats.) However, on working out the b/d parameter from different pairs of reflections at 78 °K, it was found that the internal consistency of the data itself, using Ag $K\alpha$ radiation, yielded a better and fairly precise indication of the effective temperature. If one assumes $\Theta=121$ °K, the scatter in the results of the different pairs is less than if one assumes $\Theta=95\cdot9$ °K. Plotting the data as in Figure 3, an effective characteristic temperature of $\Theta=127$ °K is indicated by linear extrapolation. The results of recalculating with $\Theta=127$ °K are also plotted in the figure. While the effective values for Θ thus obtained are satisfactory for the various orders of (111) reflections used here, they are not properly comparable with other Θ determinations. The average parameter deduced from the chosen three pairs of reflections at 78 °K assuming $\Theta = 127°$ is $b/d = 0\cdot40403 \pm 0\cdot00008$, whereas reducing Θ by 6° lowers the calculated parameter by only 0·00001.

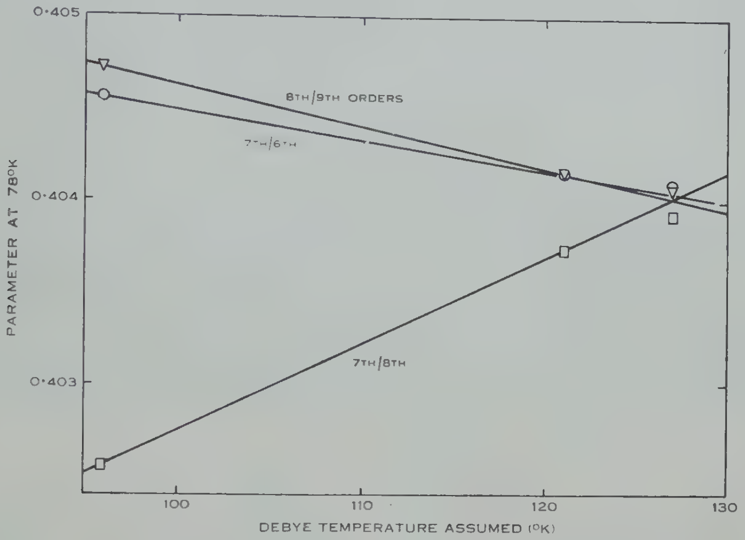


Fig. 3.—Curves determining the effective characteristic temperature at 78 °K from self-consistency of the parameters obtained from the three pairs of Ag *K*α reflections.

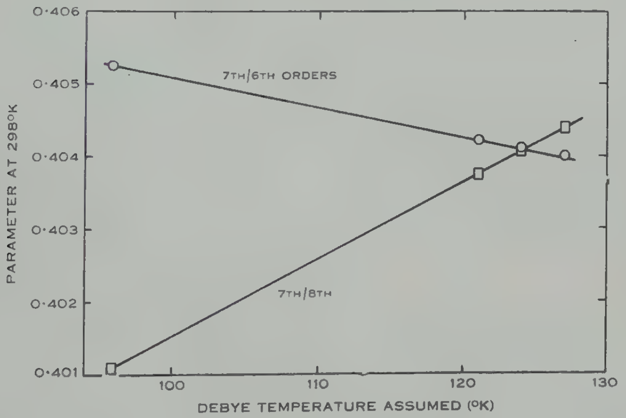


Fig. 4.—Curves for determining the effective characteristic temperature at room temperature from two pairs of Ag *K*α reflections.

The room temperature data were treated similarly and are plotted in Figure 4. Only two pairs of reflections were available here, and the intensities were less accurate, but interpolation indicated 124 °K as the effective Debye temperature. Calculating on this basis, $b/d = 0\cdot40408 \pm 0\cdot00047$, whereas on the basis of

Θ=127 °K the parameter is 0·00011 higher than this and on the basis of Θ=121 °K it is 0·00011 smaller than this.

The temperature factor at 4·2 °K is almost uninfluenced by the assumed magnitude of Θ, for here nearly the entire factor is the zero-point energy contribution. Accordingly, it was not profitable to attempt to evaluate Θ by a self-consistency method; if one assumes Θ=95·9°, the average of the three pairs of Ag $K\alpha$ reflections gives b/d=0·40436; with Θ=107°, b/d is again 0·40436, and with Θ=127°, b/d=0·40437. The internal consistency is slightly better with the latter assumption than with the others, so this appears the better value.

There remains some uncertainty in the proper magnitude of the zero-point contribution, but again the use of three pairs of reflections averages out quite effectively the errors arising from this source to an extent that probably makes them smaller than errors in the original integrated intensity measurements. For example, if the logarithms of the ratios of integrated intensities for the three pairs of reflections are corrected by adding Debye-Waller $\Delta(2M \log_{10} e)$ terms that are arbitrarily taken as only two-thirds as large as those calculated for Θ=127 °K, the resultant parameter calculated from the average of the three pairs is only increased by 0·00002.

THERMOELECTRIC POWER OF DILUTE COPPER AND SILVER ALLOYS*

By F. J. Blatt,†‡ M. Garber,† R. H. Kropschot,†§ and B. Scott†

[*Manuscript received September* 25, 1959]

Summary

Measurements of the thermoelectric power of several dilute copper and silver alloys are reported. The thermoelectric powers of the alloys were measured between about 8 and 350 °K. The results have been analysed using the Mott-Friedel relation. We have found that good agreement between the calculated and measured thermoelectric powers can be obtained only if it is assumed that the absolute thermoelectric power of pure copper and pure silver consists of two contributions: one the usual free-electron contribution and another, which is believed to be due to phonon drag. This phonon drag contribution is of the correct magnitude and shows the expected temperature dependence. Moreover, the effect of alloying on the phonon drag contribution is in agreement with theoretical estimates.

I. Introduction

Although there are numerous experiments which yield better information on the band structure of pure metals than do transport measurements, nearly all of these more sensitive procedures (Bommel 1955; Fawcett 1956; Kip *et al.* 1957; Pippard 1957*a*, 1957*b*; Harrison 1958; McClure 1958; Galt *et al.* 1959; Morse and Gavenda 1959) are inapplicable to alloys because they require that the relaxation time of electrons be relatively long. In alloys in which the residual resistivity is comparable to the ideal resistivity at room temperature such resonance techniques as discussed in the references listed above can be expected to succeed only if the experiments could be performed with success on the pure metal at room temperature. To date this has not proved feasible, and it is for this reason that transport measurements provide much of the existing information on the band structure of alloys. In addition to transport effects, equilibrium properties such as magnetic susceptibility (Hedgcock 1956; Henry and Rogers 1956) and specific heat (Rayne 1956, 1958) also shed much light (or confusion) on the alloy problem.

Until recently most theoretical work on the properties of dilute alloys was based on the rigid band model (Friedel 1954). During the past year, however, it has been suggested that in some instances, particularly in the case of certain monovalent metal alloys, this model may be totally inadequate and that the shape of the conduction band of a dilute alloy may differ significantly from the conduction band of the pure solvent metal, especially in the crucial region near the Fermi energy (Cohen and Heine 1958). If, indeed, the rigid band model fails we would expect that many of the predictions based on it should also prove incorrect.

---

* Supported in part by the Office of Ordnance Research, U.S. Army.

† Department of Physics, Michigan State University, U.S.A.

‡ Present address: Clarendon Laboratory, Oxford University, England.

§ Present address: Cryogenics Laboratory, National Bureau of Standards, Boulder, Colorado, U.S.A.

The situation at the present time is somewhat confused. The electronic specific heat of copper alloys supports the non-rigid band model of Cohen and Heine. On the other hand, measurements of the magnetic susceptibility of copper and silver alloys are most naturally explained in terms of a rigid band model. The linear dependence of the residual resistivities of copper and silver alloys on solute concentration (Linde 1931, 1932*a*, 1932*b*) also may lend support to the rigid band scheme. Yet, Ziman (1959*b*) has based a rather plausible explanation of the anomalous resistivity minimum of many very dilute copper alloys on the work of Cohen and Heine.

The measurement of the thermoelectric powers (TEP) of dilute noble metal alloys is a rather natural extension of the work of Linde (1931, 1932*a*, 1932*b*) on the residual resistivities of these alloys. Since the greatest value of that work was derived from the systematic manner in which this study was carried out we decided to investigate similarly the TEP of the same group of alloys. Ultimately this work may shed some light on the rigid band–modified band arguments ; the thermoelectric power is rather sensitive to the shape of the energy band near the Fermi energy. At this time the data are still too sparse and our understanding of the TEP of pure metals is still too rudimentary to permit us to apply our results in this manner. Nevertheless, we are able to draw some conclusions which we believe to be of interest.

This paper then constitutes in effect an interim progress report of a rather extensive study of the thermoelectric properties of dilute noble metal alloys. The current research programme and the work planned for the immediate future will be summarized at the end of this article.

To date we have made measurements on well-annealed samples of the following alloys : CuZn, CuGe, CuCd, CuIn, CuSn, CuSb ; AgZn, AgGe, AgCd, AgIn, AgSn, AgSb. With the possible exception of the cadmium and zinc alloys the solute concentration was one atomic per cent. The zinc and cadmium concentrations may have been slightly less than this amount due to loss by evaporation during melting and annealing.

We have measured the difference between the TEP of the alloys and of the pure solvent metals, copper and silver. These measurements were made on annealed alloy versus pure annealed solvent thermocouples. The absolute TEP of each solvent metal was determined by measuring the thermal e.m.f. of a solvent metal versus pure lead thermocouple (Christian *et al.* 1958). The method of alloy preparation has been discussed in some detail elsewhere (Blatt and Kropschot 1960).

II. Thermoelectric Power

The absolute TEP of a free-electron gas is given by (Wilson 1954)

$$S=\frac{\pi^2k^2T}{3e}\left(\frac{\partial \ln \sigma}{\partial \varepsilon}\right)_\eta=-\frac{\pi^2k^2T}{3e}\left(\frac{\partial \ln \rho}{\partial \varepsilon}\right)_\eta. \qquad \ldots\ldots \ (2.1)$$

Here S is the absolute TEP, k is Boltzmann's constant, T is the temperature, e is the electronic charge, σ and ρ are the electrical conductivity and resistivity respectively, ε is the electronic energy, and η is the Fermi energy. If the effective mass m^* and the relaxation time τ are assumed isotropic, the expression for the conductivity is

$$\sigma=-\frac{4e^2}{3m^*}\int_0^\infty N(\varepsilon)\tau(\varepsilon)\varepsilon\frac{\partial f_0}{\partial\varepsilon}d\varepsilon, \qquad \ldots\ldots\ldots\ldots (2.2)$$

where f_0 is the Fermi distribution.

The resistivity of an alloy is the sum of the ideal resistivity ρ_0, and of the residual resistivity ρ_r (Matthiessen's rule). Consequently, the difference between the TEP of an alloy and of the pure solvent metal is given by

$$\Delta S=-\frac{\pi^2k^2T}{3e}\left[\frac{\partial\ln(\rho_0+\rho_r)}{\partial\varepsilon}-\frac{\partial\ln\rho_0}{\partial\varepsilon}\right]_\eta. \qquad \ldots\ldots (2.3)$$

Although we have assumed in writing equation (**3.3**) that the Fermi energy of the alloy and of the pure metal are the same, it is a simple matter to account for the change in the Fermi energy due to alloying provided the rigid band model is used. One then finds that for concentrations of the order of one atomic per cent. the error introduced by the neglect of the change in Fermi energy is negligibly small (see footnote in Domenicali (1958)).

Simple algebraic manipulation leads from equation (**2.3**) to the more convenient expression first given by Friedel (1954)

$$\Delta S=-S(1-\Delta x/x)/(1+\rho_0/\rho_r), \qquad \ldots\ldots\ldots\ldots\ldots (2.4)$$

where

$$x=\eta(\partial\ln\rho_0/\partial\varepsilon)_\eta,\ \Delta x=\eta(\partial\ln\rho_r/\partial\varepsilon)_\eta. \qquad \ldots\ldots (2.5)$$

The only parameter which appears in equation (**2.4**) and which cannot be determined with relative ease is Δx. The quantity x is simply related to the absolute TEP of the pure metal; in fact, it is given by $x=288(S/T)$, if S is in μV/degK. The resistivity of the pure metal is known from the literature, and the residual resistivity of the alloy is readily determined. The parameter Δx could in principle be calculated from a knowledge of the scattering potential which one associates with a given substitutional impurity. Some information regarding this potential is known since it must satisfy certain self-consistency requirements and must also lead to the correct residual resistivity (Abelès 1953; Blatt 1955*a*; Roth 1956). Unfortunately, a variety of potentials of considerably different shapes are known to lead to practically identical residual resistivities even though the dependence of the scattering cross section on electron energy may vary significantly among these diverse potentials (Blatt 1955*b*; Roth 1956). In the following we shall consider Δx as an adjustable parameter whose value will be determined from the experimental results. This procedure does not, however, permit us as much freedom in fitting our data as may at first appear. It is evident that Δx must satisfy the following conditions:

1. Δx is independent of temperature,
2. Δx is independent of solute concentration for small concentrations.

Since the results to be reported here are on alloys of only one solute concentration for each solute element, the second requirement is of no immediate concern. However, the absence of any temperature dependence is of import to us, for our data extend over a wide range of temperature in which the TEP of the pure solvent metals show significant and somewhat unusual variations.

The thermoelectric e.m.f.'s were measured by standard techniques, using a sensitive d.c. microvolt amplifier and a microvolt potentiometer. The temperature difference between the hot and cold junctions was determined by means of calibrated copper versus constantan and copper versus gold+2·1% cobalt thermocouples.*

The range between 4·2 °K and about 350 °K was spanned in three intervals, with data from the lower intervals overlapping the higher ones ; the cold junction temperatures were maintained at either 4·2 °K (immersed in liquid helium), 79·3 °K (immersed in liquid air), or 273·2 °K (immersed in distilled water and ice). Data were taken on both heating and cooling of the hot junction, and the results were found to agree well with each other. For each sample thermocouple enough data was obtained so that at least 150 points, approximately equally spaced, fall into the entire temperature interval.

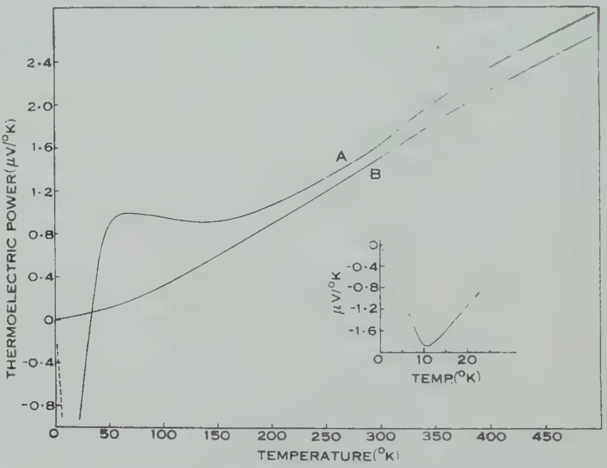


Fig. 1.—Curve *A*, the absolute thermoelectric power of pure annealed copper, measured. Curve *B*, the Sondheimer-Wilson interpolation formula, matched to curve *A* at 1173 °K.

The absolute TEP of our pure annealed copper is shown in Figure 1. This information was obtained by measuring the TEP of a pure copper versus pure lead thermocouple and subtracting therefrom the absolute TEP of pure lead (Christian *et al.* 1958).

It is apparent that the TEP of pure copper exhibits certain anomalies in the low temperature region. First there is the maximum near 70 °K which has been reported previously (Borelius, Keesom, and Johannson 1928 ; Borelius *et al.* 1930, 1932) and appears to be well established. However, we find, contrary to previous results (MacDonald 1953; MacDonald and Pearson 1953), that the absolute TEP of pure copper changes sign near 30 °K and exhibits a pronounced minimum of unusual magnitude near 10 °K. The behaviour is similar to that shown by very dilute copper alloys investigated by MacDonald (1953) and MacDonald and Pearson (1953, 1955), and at first suggested to us that our "pure" copper might not be of as high a purity as we would have wished. Although this possibility cannot be excluded entirely, we do feel that curve *A* of Figure 1 is indeed characteristic of pure copper for the following reasons.

---

* We are grateful to Drs. R. L. Powell and M. D. Bunch of the Cryogenics Laboratory, National Bureau of Standards, for providing us with these calibrated thermocouple wires.

1. The measured resistance ratio $R_{4\cdot2}/(R_{273}-R_{4\cdot2})$ is very small ($\sim 2\times10^{-3}$). Consequently it is unlikely that the material contained impurities in significant amounts.

2. Recently Powell, Roder, and Hall (1959) examined the TEP of pure annealed copper versus pure cold-worked copper thermocouples. The metal used was of extremely high purity, and our results are in satisfactory agreement with theirs, especially at low temperatures (Kropschot and Blatt 1959).

3. We have observed the same behaviour in pure silver, which until recently was also thought to have a positive absolute TEP at all temperatures. During the past year, however, measurements at Ottawa have shown that the absolute TEP of both pure silver and pure gold are negative at low temperatures and are of unusually large magnitude (MacDonald, Pearson, and Templeton 1958).

4. We have been informed by Dr. Pearson (personal communication) that he also has obtained essentially identical data for several of his pure copper samples.

The TEP's of our alloy versus pure copper thermocouples are shown in Figures 2 and 3. These TEP's are, of course, just the quantities ΔS for the various alloys.

The absolute TEP of pure silver, obtained exactly as that of pure copper, is shown in Figure 4, curve *A*. The TEP's of the alloy versus pure silver thermocouples are shown in Figure 5.

Figures 6 and 7 show the difference between the calculated and experimentally determined ΔS values for the copper and silver alloys. The calculated ΔS values were obtained by fitting equation (**2.4**) to the data at 320 °K in the case of the copper alloys and 360 °K for most of the silver alloys (arrows indicate fitting points), and evaluating the parameters Δx from this fit. The ΔS values at other temperatures were then calculated from equation (**2.4**) using, for each alloy, the appropriate Δx obtained from the high temperature fit. It is apparent that the discrepancy between the calculated and experimental points is considerable, especially as one proceeds to temperatures below about 150 °K. We would, from the above, be inclined to conclude that the Friedel relation is not adequate and that perhaps the fault lies with the rigid band model.

There is, however, another interpretation which one may attach to these results. We must remember that the Friedel relation, equation (**2.4**), is based on the free-electron model in the strictest sense. That is, the assumption is made that the TEP arises entirely as a result of the change in electron distribution with temperature, *the electron gas being considered an isolated system.* We know, however, that the electrons interact with lattice vibrations, this interaction being in fact the mechanism by which thermal equilibrium becomes established.

This interaction should give rise to an additional contribution to the TEP, the so-called phonon drag contribution (Gurevich 1945, 1946; Herring 1954; Klemens 1954; Sondheimer 1956). Obviously, the Friedel relation cannot be applied to this contribution since it is of entirely different origin from that which gives rise to the free-electron TEP.

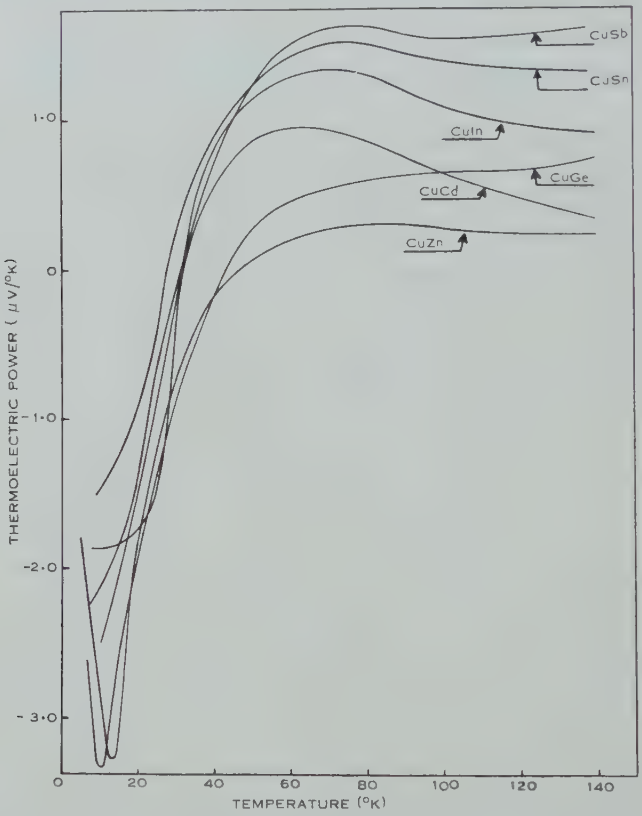


Fig. 2.—The thermoelectric power of annealed copper alloy *v.* pure annealed copper thermocouples between about 8 and 140 °K.

We therefore proceed to the following interpretation of our results. Let the total absolute TEP of pure copper be

$$S=S_e+S_g, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.6)$$

where S_e is the " electronic ", i.e. free electron, part and S_g is the phonon drag contribution. We now apply equation (**2.4**) to ΔS_e, the change in the electronic contribution only, and then compare the calculated and experimental results once more.

This latter procedure unfortunately requires one more piece of information, namely, the manner in which S is to be divided into its two components. We arbitrarily make this division by the following procedure.

The phonon drag contribution is known to decrease with increasing temperature as $1/T$ in that temperature range in which the predominant scattering mechanism of phonons is anharmonic coupling between phonons (Hanna and Sondheimer 1957). Now the absolute TEP of pure copper is known to 1173 °K (Nystrom 1948), and we assume that at this high temperature the phonon drag contribution is vanishingly small compared to S_e. The temperature dependence of the absolute TEP of pure copper above room temperature supports this contention, following closely the free-electron prediction.

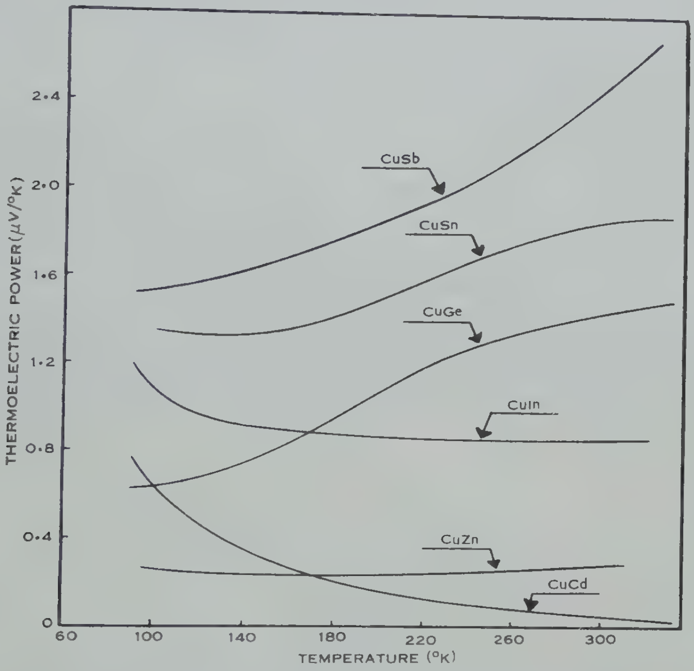


Fig. 3.—The thermoelectric power of annealed copper alloy *v.* pure annealed copper thermocouples between 100 and 340 °K.

Finally, we still proceed with some trepidation because the free-electron theory predicts the incorrect sign for the absolute TEP of copper and silver. Nevertheless, we assume that this theory does give the correct temperature dependence for S_e, and use the Sondheimer-Wilson interpolation formula (Wilson 1954), matched at 1173 °K to the TEP of copper, to determine the value of S_e at all lower temperatures. What we have done is to assume that, whatever may be the mechanism responsible for the positive sign of the TEP of copper, *this mechanism is temperature independent.* A mechanism satisfying this requirement has been proposed by Jones (1955).

The various assumptions and uncertainties inherent in this procedure cannot be justified with anything approaching rigour, and we make no attempt to do so. Instead, we merely content ourselves with the presentation of the results which we obtain in this manner, and are currently continuing experimental work along such lines that future data may bear most heavily on these ideas.

We present first the results on the copper alloys. The corresponding curves for the silver alloys are not as unequivocal as we might have hoped, and the measurements are now being repeated. In the case of the silver alloys we encountered some difficulties which are discussed below.

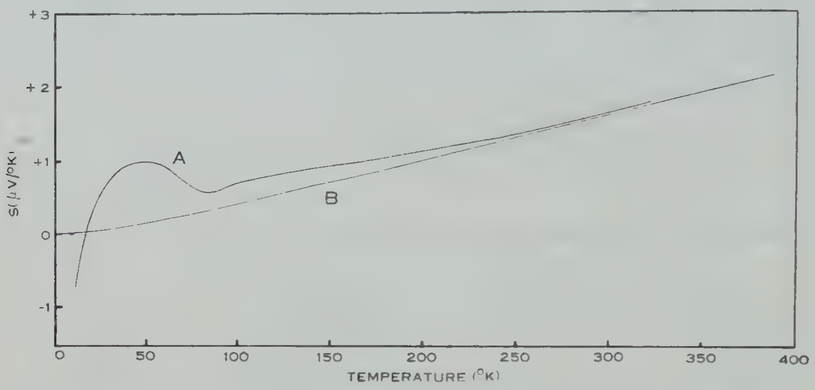


Fig. 4.—Curve *A*, the absolute thermoelectric power of pure annealed silver, measured. Curve *B*, the Sondheimer-Wilson interpolation formula (reasonable estimate).

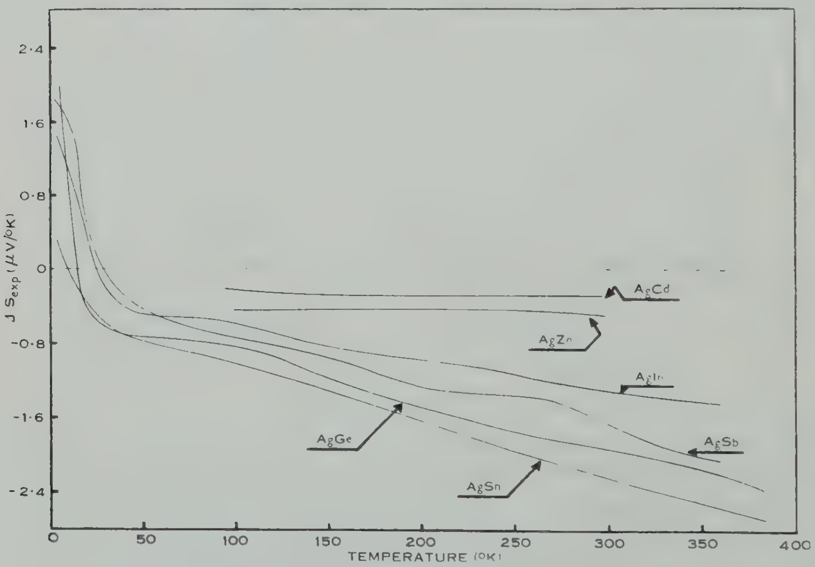


Fig. 5.—The thermoelectric power of annealed silver alloy *v.* pure annealed silver thermocouples between about 8 and 360 °K.

Curve *B*, Figure 1, shows the Sondheimer-Wilson interpolation curve fitted to the absolute TEP of copper at 1173 °K. From this curve we now obtained the values of x at each temperature. We again fitted our data to equation (**2.4**) at 320 °K, assuming that ΔS_g, the change in S_g upon alloying, is negligibly small at 320 °K. The difference between the calculated ΔS_e (from equation (**2.4**)) and the experimental values of $\Delta S = \Delta S_e + \Delta S_g$ are shown in Figure 8.

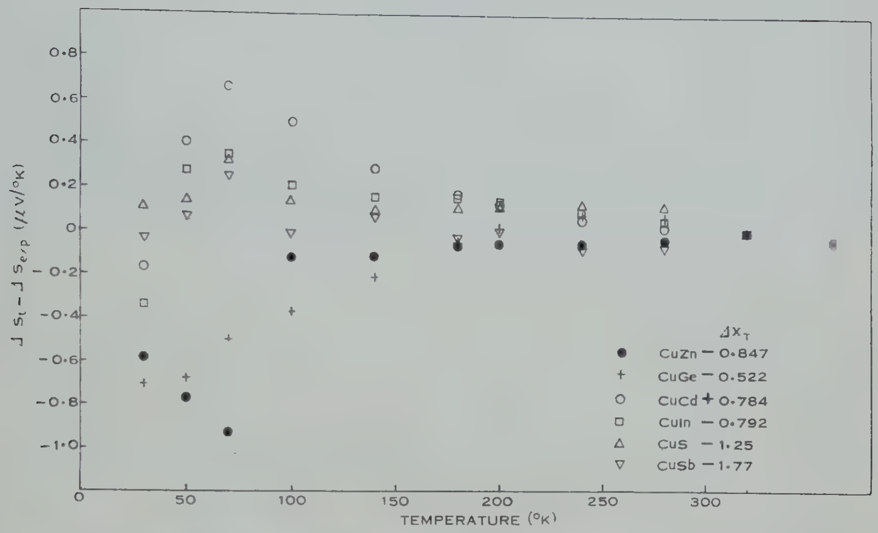


Fig. 6.—The difference between the calculated and experimental thermoelectric power of copper alloy *v.* copper thermocouples. The calculated points were obtained under the assumption of no phonon drag in either the pure metal or the alloy.

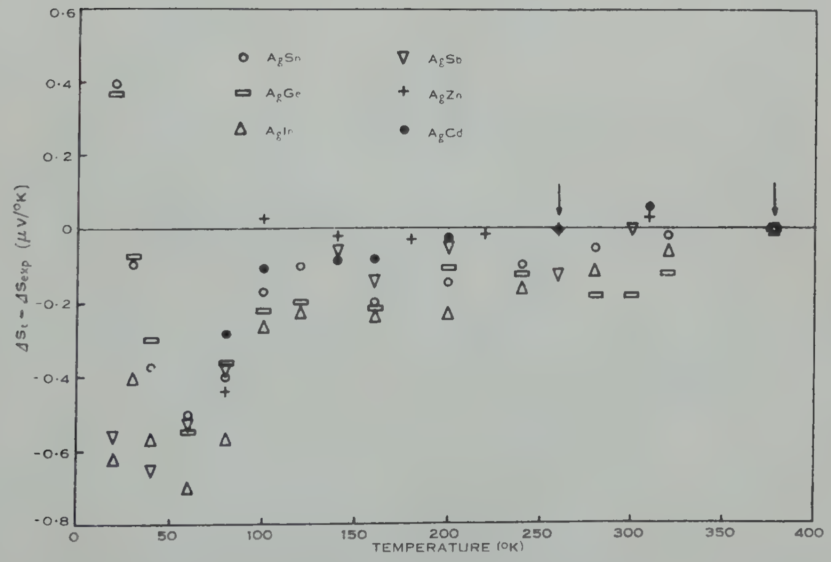


Fig. 7.—The difference between the calculated and experimental thermoelectric power of silver alloy *v.* pure silver thermocouples. The calculated points were obtained under the assumption of no phonon drag in either the pure metal or the alloy.

A pattern of alloy behaviour now emerges quite clearly. For the alloys containing zinc and germanium the Friedel theory gives very good agreement with experiment to temperatures of about 50 °K. However, alloys containing elements neighbouring silver in the periodic table (denoted as " silver-group alloys " hereafter) result in a curve which is practically the same for all the alloys and which is, moreover, almost identical to the curve obtained by taking the difference between curve *A* and curve *B* of Figure 1. This difference curve is just the assumed phonon drag contribution of pure copper, i.e. S_g^{Cu}, shown dashed in Figure 8. The solid curve of Figure 8 is the difference between the absolute TEP of copper at temperature *T* and the absolute TEP of copper at $T=320$ °K, where the fit was made.

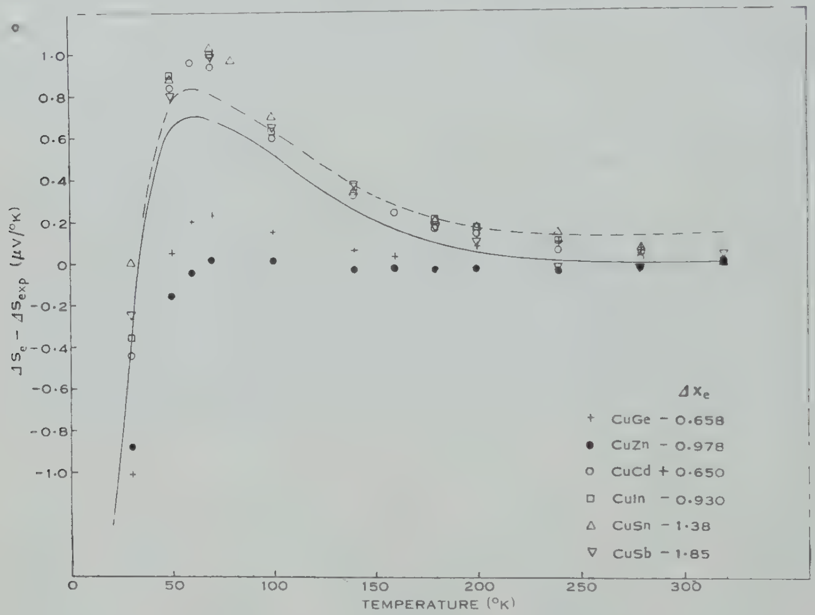


Fig. 8.—The difference between the calculated " electronic " thermoelectric power and the total measured thermoelectric power of copper alloy *v.* pure copper thermocouples. The calculations assumed that curve *B*, Figure 1, is the " electronic " contribution to the thermoelectric power of pure copper. The dashed curve here is the difference between curve *A* and curve *B* of Figure 1, ascribed to " phonon drag ".

We suggest that there is a very simple and straightforward interpretation of these results. The change in the TEP due to alloying is

$$\Delta S = \Delta S_e + \Delta S_g = S_e^{\mathrm{alloy}} - S_e^{\mathrm{Cu}} + S_g^{\mathrm{alloy}} - S_g^{\mathrm{Cu}}.$$

Now ΔS_e has been calculated by application of equation (2.4). Figure 8 then tells us the following. For the zinc and germanium alloys

$$\Delta S_e^{\mathrm{calc}} - \Delta S_{\mathrm{exp}} \simeq 0\,; \quad \text{hence } \Delta S_g \simeq 0.$$

Consequently,

$$S_g^{\mathrm{alloy}} \simeq S_g^{\mathrm{Cu}} \quad \text{(copper-group alloys)}$$

in the temperature range between 50 and 320 °K.

For the silver-group alloys we see that in the same temperature range

$$\Delta S_e^{\text{calc}} - \Delta S_{\text{exp}} \simeq S_g^{\text{Cu}}\,;$$

consequently,

$$S_g^{\text{alloy}} \simeq 0 \quad \text{(silver-group alloys)}.$$

Before proceeding we might remark that if, indeed, $S_g^{\text{alloy}}=0$ for the silver-group alloys then our analysis should be modified slightly. We have assumed in fitting the data to the Friedel relation at 320 °K that at 320 °K $\Delta S_g=0$. Therefore, what we should now do is determine ΔS_e at 320 °K by adding to $\Delta S_{\text{exp}}(320)$ of the silver-group alloys the quantity $S_g^{\text{Cu}}(320)$. We have done this, used the new value of $\Delta S_e(320)$ to determine the appropriate parameters Δx_e, and show the results in Figure 9. It is apparent that this slight modification does not produce a significant change in the curves.

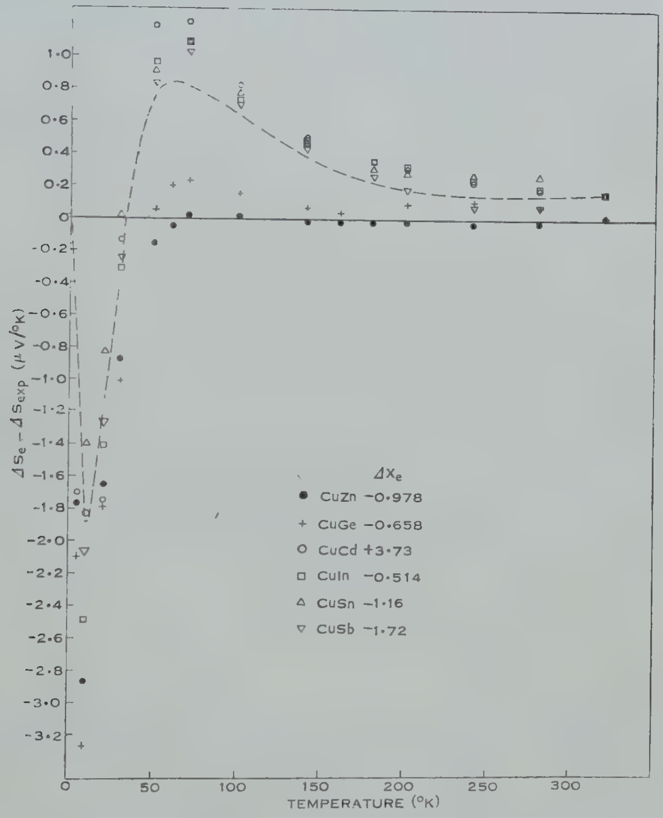


Fig. 9.—Same as Figure 8, except that here the calculation for the silver-group alloys was based on the assumption that the phonon drag contribution in these alloys vanishes at 320 °K.

Below about 50 °K all the points fall roughly along a curve which approximates the difference between curves *A* and *B* of Figure 1. Two possible explanations for this behaviour suggest themselves. First, if the low temperature anomaly of our pure copper is perhaps to be attributed to slight contamination of the pure copper, then this anomalous behaviour should be removed by the addition of further impurities such as to bring the solute concentration to one atomic per cent. Another explanation, which we believe to be the more likely in view of the earlier remarks concerning the TEP of pure copper, is that in the low temperature region the sign reversal in pure copper is a consequence of an Umklapp phonon drag mechanism (Bailyn 1958 ; Ziman 1959*a*). We also recall that at low temperatures dislocations are quite effective in scattering phonons (Klemens 1956). Now annealing is believed to reduce the dislocation density in pure copper much more than in dilute alloys (Kemp, Klemens, and Tainsh 1957). It is then rather likely that even in the annealed alloys the dislocation density may still be high enough to eliminate phonon drag at low temperatures, regardless of the type of impurity present.

In the temperature region above 50 °K the dashed curve of Figure 9 and the points for the silver-group alloys show the behaviour expected of phonon drag TEP. The decrease at higher temperatures follows approximately a $1/T$ law ; the maximum appears at a reasonable temperature and is of the right order of magnitude.

The difference between the phonon drag effect in the copper-group and silver-group alloys is most readily understood by considering the difference between copper-group and silver-group impurities as phonon scatterers. Because of the large mass difference between copper and any of the silver-group impurities the lattice thermal conductivity of the silver-group alloys is only a fraction of the lattice thermal conductivity of pure copper in the vicinity of 100 °K. Zinc and germanium, on the other hand, do not present a large cross section for scattering of phonons, and consequently the phonon drag TEP in these alloys is nearly the same as in pure copper. One can quickly arrive at order of magnitude estimates, using the results of Hanna and Sondheimer (1957), Klemens (1957), and Ambegaokar (1959), and one then finds that for our alloys

$$1\cdot0>\frac{S_g^{\text{alloy}}(\text{Cu-group})}{S_g^{\text{Cu}}}\geqslant0\cdot8\ ;\quad \frac{S_g^{\text{alloy}}(\text{Ag-group})}{S_g^{\text{Cu}}}\leqslant0\cdot16$$

in the temperature region about 100 °K. These numerical estimates are evidently in good agreement with the above conclusions concerning phonon drag in copper alloys.

The corresponding results for the alloys of silver are shown in Figure 7. Here again the Friedel relation leads to significant discrepancies as one proceeds to lower temperatures. Again, the separation of the TEP of pure silver into two components, S_e and S_g, and application of equation (**2.4**) to ΔS_e only, leads to the results shown in Figure 10. Once more we find that there is reasonably good agreement over a fairly wide temperature range, and that at lower temperatures there appears a hump suggestive of a phonon drag effect. We would have anticipated that this hump be pronounced only for the silver-zinc and silver-germanium (silver with copper-group) alloys. However, we find that also AgSn exhibits a pronounced maximum in Figure 10. None of these maxima are as large as the difference between the TEP of pure silver and the Sondheimer-Wilson curve of Figure 4 (curve A minus curve B). We are inclined to view the data on the silver alloys with some scepticism and are planning to repeat the measurements on new samples. There is reason to believe that some contamination may have occurred during sample preparation. Also we have information on the absolute TEP of silver only to about 150 °C. Consequently, we were forced to make a rather rough estimate as to a reasonable Sondheimer-Wilson fit to the TEP of silver. New alloys are being prepared by melting in vacuum in an induction furnace, the constituents being contained in a graphite crucible. Also we shall extend the measurement on the TEP of pure silver to higher temperatures by measuring the thermoelectric e.m.f. of a pure silver versus platinum thermocouple; the absolute TEP of platinum is known to high temperatures from the work of Nystrom (1948).

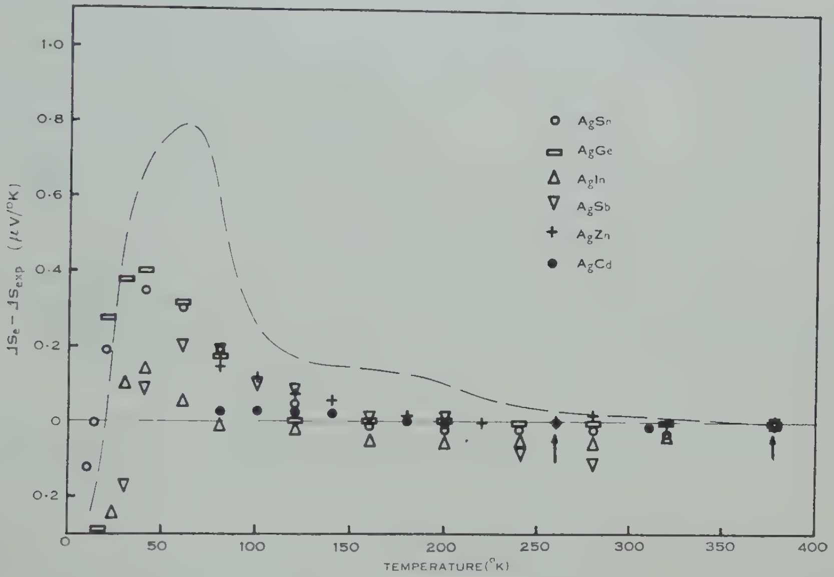


Fig. 10.—The difference between the calculated "electronic" thermoelectric power and the total measured thermoelectric power of silver alloy *v.* pure silver thermocouples. The calculations assumed that curve *B*, Figure 4, is the "electronic" contribution to the thermoelectric power of pure silver. The dashed curve here is the difference between curve *A* and curve *B* of Figure 4, ascribed to "phonon drag".

III. Conclusion and Remarks on Current Work

The measurements on the TEP of copper alloys are most readily interpreted under the assumption of a phonon drag contribution to the absolute TEP of pure copper. These results, therefore, suggest that such a contribution indeed exists in copper, and the magnitude and temperature dependence which we obtain from our data are in agreement with current theory. The data on the silver alloys are not as unequivocal as those on the copper alloys for reasons which may possibly be traced to experimental difficulties and inadequate information on the TEP of pure silver.

To further substantiate, or negate, the interpretation which we have given we are now embarked on the following experimental programme.

First, and perhaps of greatest interest, will be the measurement of the TEP of isotopically enriched copper 63. Natural copper contains both copper 63 and copper 65 in roughly 2 : 1 abundance. From the work on natural and isotopically enriched germanium (Geballe and Hull 1957) it is now firmly established that at low temperature the lattice thermal conductivity of the isotopically pure material is significantly higher than that of the element in its natural isotopic constitution. We estimate that the lattice thermal conductivity of 95 per cent. enriched copper 63 should be one-and-a-half to twice that of natural copper near 60 °K (provided both samples are equally free of impurities). Consequently, the phonon drag TEP, if it is present in natural copper, should be significantly enhanced in the isotopically enriched material.

Next, we may recall that the parameter Δx in the Friedel relation should be independent of solute concentration for sufficiently small concentrations. Measurements similar to those reported here but using alloys of various concentrations should be extremely informative. If our analysis of the present data is correct we would expect that the same values of Δx_e will give good agreement for a variety of solute concentrations. New alloys are now being prepared and the indicated measurements will be made during the coming months.

Silver and gold as impurities in copper should give rise to a significant change in the TEP, especially at low temperatures, even though these elements have the same valence as copper. In these alloys the change in the TEP should be almost entirely due to the suppression of the phonon drag TEP in the alloy.

Finally, we are initiating an investigation of dilute ternary alloys. From the data which we have and shall obtain on dilute binary alloys we expect to be able to predict with some degree of accuracy the TEP of dilute ternary alloys. For example, if one were to add to a one atomic per cent. CuCd binary alloy one atomic per cent. of indium, the change in the TEP of the alloy due to the presence of indium should reveal itself solely as an effect on what we have called the " electronic contribution " to the TEP, S_e, since the phonon drag contribution in the alloy has already been eliminated by the presence of the cadmium atoms. On the other hand the same addition of indium to a CuZn alloy should influence both S_e and S_g, since the latter is presumably still of importance in the CuZn alloy. Some studies of ternary alloys have already been reported by Domenicali (1958). Unfortunately, his results do not bear too closely on these arguments since his data encompass the temperature range extending from room temperature upward. In this temperature region the phonon drag effect is already quite small, and the influence of alloying on it would be difficult to extract from the data. It is encouraging, however, that Domenicali is able to predict with rather good reliability the TEP of the ternary alloys from his data on binary alloys.

IV. References

Abelès, F. (1953).—*C.R. Acad. Sci., Paris* **237** : 796.

Ambegaokar, V. (1959).—*Phys. Rev.* **114** : 488.

Bailyn, M. (1958).—*Phys. Rev.* **112** : 1587.

Blatt, F. J. (1955*a*).—*Phys. Rev.* **99** : 1708.

Blatt, F. J. (1955*b*).—*Phys. Rev.* **100** : 666.

Blatt, F. J., and Kropschot, R. H. (1960).—*Phys. Rev.* (In press.)

Bommel, H. E. (1955).—*Phys. Rev.* **100** : 758.

Borelius, G., Keesom, W. H., and Johannson, C. H. (1928).—*Commun. Kamerlingh Onnes Lab. Leiden* No. 196a.

Borelius, G., Keeson, W. H., Johannson, C. H., and Linde, J. O. (1930).—*Commun. Kamerlingh Onnes. Lab. Leiden* No. 206a.

Borelius, G., Keesom, W. H., Johannson, C. H., and Linde, J. O. (1932).—*Commun. Kamerlingh Onnes Lab. Leiden* No. 70a.

Christian, J. W., Jan, J. P., Pearson, W. B., and Templeton, I. M. (1958).—*Proc. Roy. Soc.* A **245** : 213.

Cohen, M. H., and Heine, V. (1958).—*Advanc. Phys.* **7** : 395.

Domencali, C. A. (1958).—*Phys. Rev.* **112** : 1863.

Fawcett, E. (1956).—*Phys. Rev.* **103** : 1582.

Friedel, J. (1954).—*Advanc. Phys.* **3** : 446.

Galt, J. K., Merritt, F. R., Yager, W. A., and Dail, H. W. (1959).—*Phys. Rev. Letters* **2** : 292.

Geballe, T. H., and Hull, G. W. (1957). —Proc. 5th Int. Conf. Low Temp. Phys. Madison 1957, No. 42–6.

Gurevich, L. (1945).—*J. Phys. U.S.S.R.* **9** : 477.

Gurevich, L. (1946).—*J. Phys. U.S.S.R.* **10** : 67.

Hanna, I. I., and Sondheimer, E. H. (1957).—*Proc. Roy. Soc.* A **239** : 247.

Harrison, M. J. (1958).—*Phys. Rev. Letters* **1** : 442.

Hedgcock, F. T. (1956).—*Phys. Rev.* **104** : 1564.

Henry, W. G., and Rogers, J. L. (1956).—*Phil. Mag.* **1** : 237.

Herring, C. (1954).—*Phys. Rev.* **96** : 1163.

Jones, H. (1955).—*Proc. Phys. Soc. Lond.* A **68** : 1191.

Kemp, W. R. G., Klemens, P. G., and Tainsh, R. J. (1957).—*Aust. J. Phys.* **10** : 454.

Kip, A. F., Langenburg, D. N., Rosenblum, B., and Wagoner, G. (1957).—*Phys. Rev.* **108** : 494.

Klemens, P. G. (1954).—*Aust. J. Phys.* **7** : 520.

Klemens, P. G. (1956).—" Handbuch der Physik." Vol. 24, pp. 198–280. (Springer-Verlag : Berlin.)

Klemens, P. G. (1957).—*Proc. Phys. Soc. Lond.* A **70** : 833.

Kropschot, R. H., and Blatt, F. J. (1959).—*Phys. Rev.* **116** : 617.

Linde, J. O. (1931).—*Ann. Phys., Lpz.* **10** : 52.

Linde, J. O. (1932*a*).—*Ann. Phys., Lpz.* **14** : 353.

Linde, J. O. (1932*b*).—*Ann. Phys., Lpz.* **15** : 219.

McClure, J. W. (1958).—*Phys. Rev.* **108** : 612.

MacDonald, D. K. C. (1953).—*Physica* **19** : 841.

MacDonald, D. K. C., and Pearson, W. B. (1953).—*Proc. Roy. Soc.* A **219** : 373.

MacDonald, D. K. C., and Pearson, W. B. (1955).—*Acta Met.* **3** : 292, 403.

MacDonald, D. K. C., Pearson, W. B., and Templeton, I. M. (1958).—*Phil. Mag.* **3** : 657.

Morse, R. W., and Gavenda, J. D. (1959).—*Phys. Rev. Letters* **2** : 250.

Nystrom, J. (1948).—*Ark. Mat. Astr. Fys.* A **34** (27) : 1.

Pippard, A. B. (1957*a*).—*Phil. Mag.* **2** : 1147.

Pippard, A. B. (1957*b*).—*Phil. Trans.* A **250** : 325.

Powell, R. L., Roder, H. M., and Hall, W. J. (1959).—*Phys. Rev.* **115** : 314.

Rayne, J. (1956).—*Aust. J. Phys.* **9** : 189.

Rayne, J. (1958).—*Phys. Rev.* **110** : 606.

Roth, L. M. (1956).—Thesis, Harvard University.

Sondheimer, E. H. (1956).—*Proc. Roy. Soc.* A **234** : 391.

Wilson, A. H. (1954).—" The Theory of Metals." 2nd Ed. (Cambridge Univ. Press.)

Ziman, J. M. (1959*a*).—*Phil. Mag.* **4** : 371.

Ziman, J. M. (1959*b*).—*Proc. Roy. Soc.* A **252** : 63.

BAND STRUCTURE OF MONOVALENT METALS AND THEIR ALLOYS

By P. G. Klemens*

[*Manuscript received September* 4, 1959]

Summary

The consequences of the Bloch theory of the conduction properties of metals can be evaluated only for a band of spherical Fermi surface, isotropic in all respects. Quantitative comparison with observations is thus possible only for the monovalent metals. It appears that even the monovalent metals do not satisfy this requirement, but that their Fermi surface departs significantly from sphericity. The information derived from the various conduction properties and the electronic specific heat is discussed, paying attention to Umklapp processes and phonon drag effects. The thermoelectric power is difficult to interpret. Systematic measurements of the changes of various physical properties on alloying may provide useful information.

I. Introduction

The Bloch theory of the conduction properties of metals is based on the following concepts.

(i) The electrons are an assembly of independent particles obeying Bose-Einstein statistics.

(ii) An electron moves in the periodic lattice potential without being scattered; the potential modulates the wave-function and modifies the relation between electron energy E and wave-vector $\mathbf{k}$.

(iii) The lattice vibrations can be resolved into lattice waves.

(iv) Electrons are scattered by disturbances in the lattice periodicity; these processes can be treated by perturbation theory. Static imperfections cause elastic scattering, and thermal vibrations scatter with the emission or absorption of a phonon.

The conduction properties of metals should be explicable, at least in principle, on this basis, with the following exceptions: (*a*) superconductivity seems to be beyond the scope of a single-particle treatment; (*b*) electron-electron scattering may be important in the case of some transition metals; (*c*) it is not clear whether certain low temperature anomalies, apparently connected with specific impurities, can be explained in terms of the Bloch theory.

These exceptions still leave a wide range of properties, which should be explained in terms of the Bloch theory. The question arises whether these properties would fit the theory quantitatively, if only we could evaluate mathematically the consequences of the theory, or whether there would be discrepancies.

---

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.; present address: Westinghouse Research Laboratories, Pittsburgh, Pennsylvania, U.S.A.

The major obstacle in such a calculation arises from the complexity of the electronic band structure of most metals, the Fermi surface extending over more than one zone. For this reason we are specially interested in the monovalent metals and their alloys, where the Fermi surface lies completely in one zone.

A full consideration of the Bloch theory must include two phenomena which are implied in the theory, although they were not explicitly considered in the original formulation (Bloch 1929). These are Umklapp processes and phonon drag effects, which were first considered by Peierls (1930*a*, 1930*b*, 1932).

II. The Bloch Theory

In a perfect lattice, the wave-function of an electron is a modulated plane wave $\mathbf{k}$, of energy $E(\mathbf{k})$, and the deviations from periodicity arising from the lattice vibrations cause interactions such that an electron $\mathbf{k}$ absorbs a phonon of wave-vector $\mathbf{q}$ and frequency ω to become an electron $\mathbf{k}'$ (or vice versa). In these processes, energy is conserved

$$E+\hbar\omega=E', \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

and in addition there is an interference condition

$$\mathbf{k}+\mathbf{q}=\mathbf{k}'+\mathbf{b}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

where $\mathbf{b}$ is an inverse lattice vector. In particular, processes for which $\mathbf{b}=0$ conserve the total wave-vector (or quasi-momentum) of the electron-phonon system and are the normal or N-processes. There are, however, additional processes for which $\mathbf{b}\neq 0$; these are called Umklapp or U-processes (flip-over processes).

The probability of these interactions is proportional to C^2, where the interaction parameter C is a measure of the interaction Hamiltonian. In general C depends upon $\mathbf{k}$, upon $\mathbf{q}$, and on the polarization of the lattice wave $\mathbf{q}$. For low values of q, C is independent of the magnitude q. Since C depends upon the wave-functions of the electrons in states $\mathbf{k}$ and $\mathbf{k}'$, and on the change in lattice potential due to a wave $\mathbf{q}$, and since our knowledge of these quantities is very imperfect, C can be estimated only very roughly (e.g. Bardeen 1956).

If we knew the energy contours in $\mathbf{k}$-space (in particular the Fermi surface and its derivatives) and also all the transition probabilities, we could in principle calculate all the conduction properties by means of the transport equation

$$\frac{e}{\hbar}\mathbf{F}\cdot\frac{\partial f}{\partial \mathbf{k}}+\frac{1}{\hbar}\,\frac{\partial E}{\partial \mathbf{k}}\cdot\operatorname{grad} T\frac{\mathrm{d}f}{\mathrm{d}T}=\left.\frac{\partial f}{\partial t}\right], \qquad \ldots\ldots\ldots\ldots \quad (3)$$

where f is the occupation probability of the state $\mathbf{k}$, $\mathbf{F}$ the electric field, e the electronic charge, and the right-hand side denotes the rate of change of f due to interaction processes. We may write formally

$$\left.\frac{\partial f}{\partial t}\right]=\frac{f^0-f}{\tau}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

where f^0 is the equilibrium value of f; this defines a relaxation time τ. However, $\mathrm{d}f/\mathrm{d}t]$ consists of a sum of terms of the form

$$(1-f)f'(N+1)-f(1-f')N, \qquad \ldots\ldots\ldots\ldots \quad (5)$$

where each term corresponds to the contribution from the interaction $\mathbf{k}+\mathbf{q}\longleftrightarrow\mathbf{k}'$, f' and N being the occupations of the electron state $\mathbf{k}'$ and lattice mode $\mathbf{q}$. Thus τ depends upon the deviations from equilibrium of all interacting states $\mathbf{k}'$ and modes $\mathbf{q}$, and the effective relaxation time depends upon the relative values of F and grad T. In the original formulation of the theory, the lattice modes are assumed to be in equilibrium ($N=N^0$).

The transport equation has been solved only in the case of complete spherical symmetry, i.e.

(*a*) energy contours in **k**-space are spherical,

(*b*) apart from the factor (**5**), the scattering probabilities depend only on the magnitude of q.

Such a spherically symmetrical model can apply, at best, only to the monovalent metals. However, if Umklapp processes are important, the scattering probability and $\tau(\mathbf{k})$ lose their spherical symmetry, even if the energy contours are spherical, and depend also on $\mathbf{k}$.

It can be shown quite generally, irrespective of the shape of the energy contours and their position relative to the zone boundaries, that at high temperatures the electrical resistivity $\rho(T)$ increases linearly with T. Also at high temperatures the effective value of $\tau(\mathbf{k})$ is the same for electrical and thermal conduction, so that the Wiedemann-Franz law is satisfied

$$\rho/WT=L, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{6})$$

where W is the thermal resistivity and L the Lorenz number. More generally, (**6**) holds provided only that the electron gas is highly degenerate and that the scattering processes are elastic.

Because of the difficulty of estimating C^2, it is not possible to calculate the magnitude of the resistivities, but one can derive relations between them. In the case of complete spherical symmetry, i.e. considering N-processes only, and assuming $C(q)$ to be independent of q, we obtain the following relations between $\rho(T)$ and $W(T)$ at $T_1\ll\theta$ and $T_2>\theta$, where θ is the Debye temperature (e.g. Klemens 1956):

$$\rho_i(T_1)=497\cdot 6\rho_i(T_2)[T_1^5/\theta^4T_2], \qquad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{7})$$

$$W_i(T_1)=64\cdot 0N_a^{\frac{2}{3}}W_i(T_2)[T_1/\theta]^2, \qquad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{8})$$

$$W_i(T_1)=(64\cdot 0/497\cdot 6)N_a^{\frac{2}{3}}\rho_i(T_1)[\theta^2/LT_1^3]. \qquad \ldots\ldots\ldots \quad (\mathbf{9})$$

Here N_a is the number of electrons per atom, a measure of the volume enclosed by the Fermi surface and independent of the functional form of $E(k)$. The suffix i denotes that these resistivities are " ideal " (i.e. due to the lattice waves only); a residual resistivity (ρ_0 or W_0) due to elastic scattering by imperfections must be added.

Of relations (**7**), (**8**), and (**9**), any two are independent; (**7**) and (**8**) have to be modified if $C(q)$ or the wave velocity ω/q varies with q, but (**9**), relating two low temperature resistivities, holds irrespective of such possible variations, unless these are particularly violent.

In any case, (**7**) and (**8**) are likely to fail because they do not take account of Umklapp processes, which are certain to be significant in the case of high temperatures (large values of q). At low temperatures, however, they would not occur, and thus (**9**) should be satisfied even when (**7**) and (**8**) are not.

We must now consider the proper value of θ to be used in (**7**) to (**9**); this depends upon the ratio $C_L : C_T$, where C_L is the coupling constant between electrons and longitudinal phonons, C_T similarly with transverse phonons. There are two extreme possibilities :

$$C_L = C, \ C_T = 0 \quad \text{(Bloch coupling scheme),} \quad \ldots\ldots \ (\mathbf{10a})$$

$$C_L = C_T = C/\sqrt{3} \quad \text{(Makinson coupling scheme),} \quad \ldots\ldots \ (\mathbf{10b})$$

while in general

$$C^2 = C_L^2 + 2C_T^2. \quad \ldots\ldots \ (\mathbf{11})$$

There are two Debye temperatures, θ_L and θ_T, pertaining to the two polarization branches, and $\theta_L \sim 1 \cdot 5\theta_T$. The specific heat Debye temperature is θ_D, and $\theta_D \sim \theta_T$. In the case of Bloch coupling (**10a**), the value of θ in (**7**) to (**9**) is equal to θ_L, in the case (**10b**) it is approximately equal to θ_D. In any case it cannot be less than θ_D.

III. The Fermi Surface of Monovalent Metals

When considering the observed conduction properties of the monovalent metals, the relation of greatest interest is (**9**), being independent of the complications at high temperatures. However, for all monovalent metals, the observed values of $\rho_i(T_1)/W_i(T_1)$ exceed the values calculated from (**9**) by substantial factors (e.g. Klemens 1956). This discrepancy is minimized by choosing the lowest possible value of θ, i.e. θ_D. This would imply the Makinson coupling scheme, and this is also borne out from measurements of the lattice thermal conductivity (e.g. Kemp and Klemens 1960), but even then the observed values of ρ_i/W_i are too large by factors of order 5 for the noble metals, by a factor 2 to 3 for sodium, and by factors from 10 to 15 for the other alkali metals. Quite clearly, the simple spherical model is inadequate, though sodium comes closer to it than any of the others.

A qualitative explanation of this discrepancy may be given along the following lines.

The formal solution of the transport equation is of the form :

$$f - f^0 \propto (\mathbf{k} \cdot \mathbf{F})\tau \frac{\mathrm{d}f^0}{\mathrm{d}\varepsilon} \quad \text{(electrical conduction),} \quad \ldots \ (\mathbf{12})$$

$$f - f^0 \propto (\mathbf{k} \cdot \text{grad } T)\tau\varepsilon \frac{\mathrm{d}f^0}{\mathrm{d}\varepsilon} \quad \text{(thermal conduction),} \quad \ldots\ldots \ (\mathbf{13})$$

where $\varepsilon = (E - \xi)/KT$ is the reduced energy. At low temperatures the resistive process is thus a random walk over large distances on the Fermi surface in the case of electrical conduction, but a single-step process in the case of thermal conduction. In thermal conduction, $\tau(\mathbf{k})$ depends only on the local properties of the Fermi surface, but in the case of electrical conduction it depends upon the overall shape of the Fermi surface—hence the factor $N_a^{\frac{2}{3}}$ in (**8**) and (**9**).

U-processes, if they occur at all at low temperatures, will mainly link states at the point of closest approach of the Fermi surface to the zone boundary. In the case of electrical conduction, they will affect the random walk on the Fermi surface, acting as a short cut from one point to an opposite point, and thus increase the electrical resistance. Within rather wide limits, the effect of U-processes on $\tau(\mathbf{k})$ for electrical conduction should be independent of the probability of the U-processes themselves ; $\tau(\mathbf{k})$ is still governed by N-processes, but the effective topology of the Fermi surface is altered.* Thus ρ_i still varies as T^5. On the other hand, $\tau(\mathbf{k})$ for thermal conduction is unaffected by U-processes over most of the Fermi surface ; thus ρ_i/W_i should be increased by some factor independent of T. In the case of face-centred cubic metals (group Ib), with 8 points of nearest approach, this factor seems to be of order 4 to 5 ; in the case of body-centred metals (group Ia), with 12 such points, it should be larger, and seems to be of order 10 to 15.

In the limit of lowest temperatures, U-processes occur only if the Fermi surface touches the zone boundary. In practice we observe ρ_i only above a finite temperature, so that U-processes occur because

(*a*) the Fermi surface touches the zone boundary, or

(*b*) the Fermi surface almost touches the zone boundary, or

(*c*) the Fermi surface is almost spherical, but U-processes have a high probability because the required phonons $\mathbf{q}$ have a surprisingly low velocity and/or the matrix elements for U-processes are unusually high.

In any case, it is clear that U-processes do occur at liquid helium temperatures, except possibly in the case of sodium.

The discrepancy from (**8**), i.e. the fact that $W_i(T_2)/W_i(T_1)$ is too high, is due to the enhancement of $W_i(T_2)$ partly by U-processes and partly by the dispersion of the lattice waves, while $W_i(T_1)$ is not affected by U-processes. The relatively better fit of the Bloch-Grüneisen relation (**7**) is fortuitous : both $\rho_i(T_1)$ and $\rho_i(T_2)$ are increased by U-processes.

The importance of U-processes in the low temperature electrical resistivity is also borne out by the absence of an observed phonon drag effect. In the absence of U-processes the total wave-vector of the combined assembly of electrons and phonons is conserved (**2**), and the factor (**5**) vanishes if f and N have the following quasi-equilibrium form :

$$N(\mathbf{q})=N^0(\mathbf{q})+\mathbf{d}\,.\,\mathbf{q}(\mathrm{d}N^0/\mathrm{d}x), \qquad \ldots\ldots\ldots\ldots \quad (\mathbf{14})$$

$$f(\mathbf{k})=f^0(\mathbf{k})+\mathbf{d}\,.\,\mathbf{k}(\mathrm{d}f^0/\mathrm{d}\varepsilon), \qquad \ldots\ldots\ldots\ldots \quad (\mathbf{15})$$

where $x=\hbar\omega/KT$ and $\mathbf{d}$ is some constant (Peierls 1930, 1932). An electrical field $\mathbf{F}$ tends to increase $\mathbf{d}$, but a steady state—and hence a finite value of ρ_i—can be attained only if phonons are scattered by crystal imperfections or by three-phonon U-processes. If τ' is the phonon relaxation time for such scattering processes, and if τ_0 is the phonon relaxation time for scattering by conduction electrons, then the ideal resistivity becomes

---

* Other authors (Bailyn and Brooks 1956, Pfennig 1959) have treated the effect of U-processes as being additive to those of N-processes. This may be a reasonable approximation at higher temperatures, when nearly every point of the Fermi surface can describe an electron state capable of participating in a U-process, but fails completely at low temperatures where the U-processes occur only near special points, which in turn do not contribute to the current.

$$\rho_i = \rho_B\{\tau_0/(\tau_0+\tau')\}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \mathbf{(16)}$$

where ρ_B is the ideal resistivity calculated by substituting N^0 for N into (**5**), i.e. by assuming the phonon gas to be in equilibrium.

We know enough about τ_0 and τ' from experiments on the lattice thermal conductivity of dilute alloys (e.g. Klemens 1958) to be able to say that $\tau_0/(\tau_0+\tau') \ll 1$ in well-annealed metals below 10 °K. Since $\rho_i \propto T^5$ and since ρ_i is certainly not smaller than predicted by theory, it follows that the wave-vector is not conserved in those electron-phonon interactions which contribute most to the electrical resistance. Thus, when an electron wanders from one side of the Fermi surface to the other, it will suffer a U-process somewhere along its path in most cases (80 per cent. in face-centred, more than 90 per cent. in body-centred cubic metals).

A possible exception is sodium, where the discrepancy in the ratio ρ_i/W_i is only between 2 and 3 and where at lowest temperatures $\rho_i \propto T^6$. A complicating feature is that below about 50 °K it consists of a mixture of body-centred cubic and hexagonal close-packed material. If the Fermi surface were to touch the zone boundary in the hexagonal close-packed but not in the body-centred cubic phase, then the discrepancy in ρ_i/W_i would depend upon the relative volume of the two phases. The observed discrepancy is certainly inconsistent either with strong U-processes or absence of U-processes in a single phase. The T^6 variation could imply either that U-processes are not so important in the body-centred cubic phase, so that phonon drag (**16**) becomes noticeable, or that U-processes are being gradually frozen out in that temperature range. It does not seem possible at present to distinguish between these alternatives ; in any case it confirms our picture that sodium is the one metal where U-processes are getting rare at liquid helium temperatures (at least in the body-centred cubic phase), if they are not completely absent, so that the Fermi surface certainly does not touch the zone boundary.

In the case of copper we now have good evidence, from measurements by Pippard (1957) of the high frequency skin resistance, that the Fermi surface is strongly distorted. If the penetration depth is much less than the electronic mean free path, the skin resistance is independent of the mean free path, but depends upon the curvature of the Fermi surface in **k**-space. Since the resistance shows a strong variation with surface orientation, the Fermi surface is highly anisotropic. From an analysis of his data, Pippard concluded that the Fermi surface comes very close to the zone boundary at the points of 8-fold symmetry and probably makes contact.

The magneto-resistance of group Ib metals in strong fields is also strongly suggestive of a high degree of anisotropy of the Fermi surface ; thus Justi and Scheffers (1938) found a strong orientational dependence in gold, while Chambers (1956), working on polycrystalline copper, silver, and gold, found a field dependence inconsistent with a simply connected Fermi surface (i.e. one which does not touch the zone boundary), unless the probability of scattering by imperfections is itself temperature dependent.

IV. THERMOELECTRIC PROPERTIES

Although the thermoelectric properties are not yet fully understood, they too should, in principle, provide information about the band structure. There are two components of thermoelectric power: (*a*) an electronic component which, if a relaxation time can be defined, should be proportional to T, and (*b*) a lattice component. The latter arises because an electric current corresponds to deviation **d** in the electron distribution from equilibrium (**15**), which will tend to set up a phonon distribution (**14**) with the same value of **d**, hence a heat current, because the electron-phonon interactions conserve the total wave-vector. In the absence of U-processes the lattice component should be negative, but U-processes would tend to reverse its sign (Bailyn 1959). The lattice component of thermoelectric power is decreased in magnitude by phonon-phonon interactions and by the scattering of phonons by imperfections; it can thus be reduced by cold work and by the introduction of impurity atoms of different mass (see Blatt *et al.* 1960).

The electronic component should be negative for free electrons, and its sign is in general the same as the sign of $-d\sigma(\xi)/d\xi$, where σ is the electrical conductivity as function of Fermi energy ξ, keeping all scattering processes the same. While alkali metals have the expected negative sign for their thermoelectric power at high temperatures, that of the noble metals is positive. Blatt has shown that the lattice component is positive at liquid oxygen temperatures, reversing its sign at lower temperatures, but that it is not big enough to fully account for the positive thermoelectric power at high temperatures.

The occurrence of Umklapp processes would certainly tend to reverse the sign of $d\sigma(\xi)/d\xi$ at high temperatures, as an expansion of the Fermi surface (increase of ξ) would tend to bring it closer to the zone boundaries, thus enhancing the frequency of U-processes and tending to reduce σ. It is not clear, however, why this trend towards sign reversal should be less effective in the alkali metals (where the Fermi surfaces enclose regions which, in extended **k**-space, form a close-packed array) than in the noble metals (where the inverse lattice structure is more open, but the Fermi surfaces approach each other more closely). However, if the Fermi surface touches the zone boundary, it must be concave in parts, though mainly convex, and processes for which the final state lies on the Fermi surface near the zone boundary will be enhanced if ξ is increased. A similar reason has been given for the positive sign by Jones (1955).

The explanation of the positive sign at high temperatures will also explain the positive sign in the residual resistance region, since the dominant scattering processes are similar in both cases.

The interpretation of the thermoelectric power is at present obscure, but we may conclude from the complexity of its behaviour that U-processes are important at high temperatures, and that the Fermi surface of the noble metals departs more markedly from sphericity than that of the alkali metals.

V. BAND STRUCTURE OF ALLOYS

It should be possible to derive some information about the band structure of the noble metals from the variation of the physical properties of their dilute alloys with solute content. In the simplest model of alloying all energy contours are kept fixed, and the only effect of alloying is to increase or decrease the electron concentration, thus expanding or contracting the Fermi surface (rigid band model). By studying the dependence of various properties on alloy concentration and noting discontinuities in behaviour, it should be possible to find out which particular contour just touches the zone boundary. However, if alloying also changes the shape of the energy contours in **k**-space appreciably, the analysis becomes more complicated.

It had been noted that the boundary of the α-phase of various alloys of copper and silver occurred at the composition corresponding to 1·4 conduction electrons per atom (Hume-Rothery rule). Mott and Jones (1936) explained this in terms of the rigid band model: if $N_a=1\cdot 4$, a spherical Fermi surface just touches the zone boundary, and this determines the limit of stability of the face-centred cubic structure. However, since there is now strong evidence that contact occurs already in the pure parent metal, this explanation needs re-examination.

Rayne (1957, 1958), in order to test the suggestion (Klemens 1954) that the Fermi surface of copper touches the zone boundary, investigated the electronic specific heat of a series of copper-zinc and copper-germanium alloys in the α-phase. The electronic specific heat, being a measure of the density of electron states, should increase slowly if the energy contours are spherical and if the rigid band model applies. Rayne found a rapid increase of the specific heat with electron concentration up to $N_a=1\cdot 05$, followed by a gradual saturation. The detailed interpretation of these results is still obscure, but they are clearly inconsistent with a spherical Fermi surface of pure copper.

Cohen and Heine (1958) have attempted to reconcile Jones's explanation of the Hume-Rothery rule with the fact that the Fermi surface in pure copper is strongly distorted. They suppose that in the pure metal it does indeed touch the zone boundary, but that alloying tends to make it more spherical, until it actually ceases to touch the zone boundary at $N_a=1\cdot 1$. Further alloying again expands the Fermi sphere, and at $N_a=1\cdot 4$ it again touches the zone boundary. This second contact determines the phase boundary in the manner suggested by Jones.

It is, however, not certain whether this is a unique way of explaining the Hume-Rothery rule. It is possible that the area of the Fermi surface keeps increasing with increasing electron concentration, even if the Fermi surface already touches the zone boundary at $N_a=1$, until it starts decreasing again at about $N_a=1\cdot 4$, and that this latter condition determines the phase boundary.

Alloying with transition metals should reduce the electron concentration and thus make the Fermi surface more spherical. Taylor and Coles (1956) have indeed noted an early reversal of sign in the thermoelectric power of silver-palladium alloys with increasing palladium content.

On the other hand, the introduction of nickel into copper seems to increase the density of states (Pugh *et al.* 1957; Pugh and Ryan 1958; Guthrie, Friedberg, and Goldman 1959), and in fact Guthrie (1959) noted an increase in the electronic specific heat of a ternary copper-nickel-zinc alloy of $N_a=1$. It is not clear whether this indicates a general failure of the rigid band model of the noble metal alloys, or whether this is an anomaly peculiar to copper-nickel alloys. In this context it should be noted that alloying, irrespective of valency, seems to reduce the lattice thermal conductivity of many alloys (e.g. Kemp and Klemens 1960). It is not clear how much of this effect is due to a change in the density of states and how much is due to a change in dislocation density.

Further systematic work is called for on the properties of alloys. The electronic specific heat should be studied on dilute silver-palladium and silver-cadmium alloys, as well as on alloys with $N_a=1$ but various residual resistivities and atomic masses. Similar work should be done on gold alloys. Other properties which require study on many alloy systems include the ideal electrical resistivity over a wide temperature range, and the thermoelectric power.

VI. References

Bailyn, M. (1959).—*Phys. Rev.* **112**: 1587.
Bailyn, M., and Brooks, H. (1956).—*Bull. Amer. Phys. Soc.* (II) **1**: 300.
Bardeen, J. (1956).—*Canad. J. Phys.* **34**: 1171.
Blatt, F. J., Garber, M., Kropschot, R. H., and Scott, B. (1960).—*Aust. J. Phys.* **13**: 223.
Bloch, F. (1929).—*Z. Phys.* **53**: 216.
Chambers, R. G. (1956).—*Proc. Roy. Soc.* A **238**: 344.
Cohen, M. H., and Heine, V. (1958).—*Advanc. Phys.* **7**: 395.
Guthrie, G. L. (1959).—*Phys. Rev.* **113**: 793.
Guthrie, G. L., Friedberg, S. A., and Goldman, J. E. (1959).—*Phys. Rev.* **113**: 45.
Jones, H. (1955).—*Proc. Phys. Soc. Lond.* A **68**: 1191.
Justi, E., and Scheffers, H. (1938).—*Phys. Z.* **39**: 105.
Kemp, W. R. G., and Klemens, P. G. (1960).—*Aust. J. Phys.* **13**: 247.
Klemens, P. G. (1954).—*Aust. J. Phys.* **7**: 70.
Klemens, P. G. (1956).—" Encyclopedia of Physics." Vol. 14, Pt. 4. (Springer: Berlin.)
Klemens, P. G. (1958).—*Solid State Physics* **7**: 1.
Mott, N. F., and Jones, H. (1936).—" The Theory of the Properties of Metals and Alloys." (Oxford Univ. Press.)
Peierls, R. (1930*a*).—*Ann. Phys., Lpz.* **4**: 121.
Peierls, R. (1930*b*).—*Ann. Phys., Lpz.* **5**: 244.
Peierls, R. (1932).—*Ann. Phys., Lpz.* **12**: 154.
Pfennig, H. (1959).—*Z. Phys.* **155**: 332.
Pippard, A. B. (1957).—*Phil. Trans.* A **250**: 325.
Pugh, E. W., Coles, B. R., Arrott, A., and Goldman, J. E. (1957).—*Phys. Rev.* **105**: 814.
Pugh, E. W., and Ryan, E. M. (1958).—*Phys. Rev.* **111**: 1038.
Rayne, J. A. (1957).—*Phys. Rev.* **108**: 22.
Rayne, J. A. (1958).—*Phys. Rev.* **110**: 606.
Taylor, J. C., and Coles, B. R. (1956).—*Phys. Rev.* **102**: 27.

THE LATTICE THERMAL CONDUCTIVITY OF ALLOYS

By W. R. G. KEMP* and P. G. KLEMENS*†

[*Manuscript received August* 25, 1959]

Summary

The lattice thermal conductivity of alloys at liquid helium temperatures yields information about the nature of the electron-phonon interaction, and about the density of dislocations. It appears that in copper, silver, and gold the electrons interact with transverse as well as with longitudinal lattice waves. The lattice conductivity also varies with alloy composition ; it is not clear whether this variation should be ascribed to changes in the electron density or to the presence of dislocations, locked into stable configurations by the solute atom. If the latter view is adopted, the density of dislocations in annealed alloys is typically of the order of 10^{10} to 10^{11} lines/cm^2.

Changes of the lattice thermal conductivity of deformed alloys were compared with stored energy release ; this suggests that the theory underestimates the scattering of phonons by dislocations by a factor of 3 to 6.

I. INTRODUCTION

It is possible to determine the lattice component $\varkappa_g$ of the thermal conductivity of alloys of monovalent metals if their solute content is high enough (i.e. if their residual resistivity ρ_0 exceeds, say, $0 \cdot 5$ $\mu\Omega$ cm), so that the electronic component $\varkappa_e$ is no longer very much larger, and so that $\varkappa_e$ can be calculated from the measured value of ρ_0. The method of separating $\varkappa_e$ and $\varkappa_g$ has been discussed by Kemp *et al.* (1956) and by Klemens (1956, 1958).

At liquid helium temperatures, in general, $\varkappa_g \propto T^2$. Theoretically there are two resistive mechanisms leading to that temperature dependence : scattering of phonons by electrons and by dislocations (e.g. Klemens 1956, 1958). At higher temperatures $\varkappa_g$ reaches a maximum and then decreases again, and at the latter temperatures (usually at liquid oxygen temperatures) the resistivity arises from the scattering of phonons by point defects and from the mutual interaction between phonons due to anharmonicities. These effects are discussed by White (1960).

Interest in the lattice thermal conductivity arose originally because it provides information about the scattering of phonons by electrons. The scattering by dislocations enters as a complicating factor. Estermann and Zimmerman (1952) first noted that in Cu-10% Ni $\varkappa_g$ was less in the cold-worked state than when annealed, though it still varied as T^2 at lowest temperatures. It was subsequently shown (Klemens 1955) that dislocations scatter phonons so as to contribute a lattice resistivity $W_D \propto T^{-2}$ to W_g $(=1/\varkappa_g)$; additional resistivity due to plastic deformation was in general found to vary as T^{-2} (e.g.

---

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

† Present address : Westinghouse Research Laboratories, Pittsburgh, Pennsylvania, U.S.A.

Kemp *et al.* 1956). Thus the resistivity component W_E due to phonon-electron scattering can be studied only on specimens sufficiently well annealed so that $W_D \ll W_E$. Other defects, such as point defects, have a different temperature variation of lattice thermal resistivity and their effect can thus be separated.

We shall present evidence which suggests that many alloy specimens contain a high density of dislocations even after annealing above the recrystallization temperature.

II. Scattering of Phonons by Electrons

The interaction between conduction electrons and phonons not only limits the electronic conduction properties, producing an ideal electrical resistivity ρ_i and an ideal thermal resistivity W_i, but it also limits the lattice thermal conductivity, producing a lattice thermal resistivity $W_E \propto T^{-2}$ (Sommerfeld and Bethe 1933; Makinson 1938). The magnitude of W_E can be expressed in terms of the magnitude of ρ_i at $T > \theta$, as was done by Makinson, but the latter quantity is affected to a substantial but uncertain degree by Umklapp processes and the dispersion of phonons. The ideal thermal resistivity W_i at $T \ll \theta$ is, however, not affected by these complications, and, since W_i and W_E at $T \ll \theta$ are limited by the same electron-phonon interaction processes, estimates of W_E in terms of W_i should be particularly reliable (Klemens 1954*a*). The quantity W_i for the pure parent metal is known from its low temperature thermal conductivity; corrections for any change of W_i due to alloying can be made from observed changes in ρ_i (Klemens 1959).

While the ratio W_E/W_i at low temperatures is independent of the strength of the electron-phonon interactions and insensitive to small deviations of the Fermi surface from sphericity, it depends rather critically on the nature of the coupling, i.e. on whether the electrons interact with transverse waves as strongly as with longitudinal waves, and whether the interaction of lattice waves of different polarizations with each other by anharmonicities is strong or weak compared with the phonon-electron scattering.

In the case of strong coupling between the polarization branches, W_E/W_i is higher by a factor 15–20 if the electrons interact with longitudinal phonons only (Bloch coupling scheme) than if they interact equally strongly with phonons of both polarizations (Makinson coupling scheme) (Klemens 1954*a*). Intermediate cases are also possible. Klemens (1956) has extended the theory to cases when coupling between the polarization branches is not strong. Coupling must become weak at sufficiently low temperatures, and deviations from the $\varkappa_g \propto T^2$ variation, found by Chari and de Nobel (1959) in silver alloys at about 2 °K, may be due to this cause.

III. Variation of Lattice Conductivity with Alloy Composition

While one cannot observe $W_g T^2$ at liquid helium temperatures for the pure metals directly, one can do so for a series of dilute alloys and extrapolate to zero solute content. By choosing two alloys, one of electron concentration above that of the pure metal and one of lower electron concentration, one should be able to find $W_g T^2$ for the pure metal by interpolation. This assumes, of course, that all specimens are well annealed ($W_D \ll W_E$), and that W_E varies in a regular manner with electron concentration. Thus Kemp *et al.* (1956) measured a series of silver-palladium and silver-cadmium alloys. Some of their values of W_gT^2 are plotted in Figure 1 as function of electron concentration (or solute content). Figure 2 shows a similar plot for some copper alloys, where unfortunately the data are less complete.* In the gold system we have data for Au-0·7% Pt and Au-1·8% Pt only (Birch *et al.* 1959), and there are data on very dilute alloys of Fe in Cu, Ag, and Au (White, Woods, and Elford 1959).

Even though the variation of W_ET^2 with composition need not be monotonic, especially as the Fermi surface of copper, silver, and gold probably touches, or almost touches, the zone boundary (Klemens 1956, Cohen and Heine 1958), it

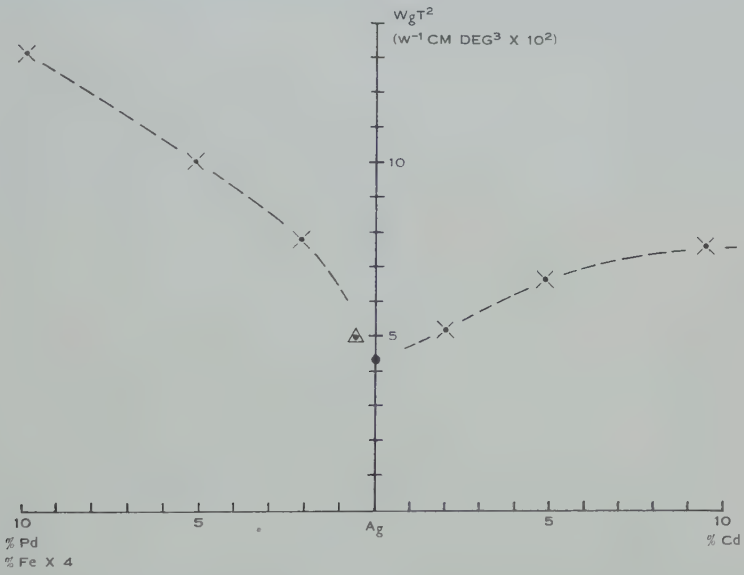


Fig. 1.—Values of W_gT^2 at liquid helium temperatures for various annealed silver alloys: ×, Ag-Pd and Ag-Cd (Kemp *et al.* 1956); △, Ag-Fe (White, Woods, and Elford 1959); ●, theoretical value for W_ET^2 (Klemens 1954*a*).

should vary regularly with electron concentration and not show the cusplike variation of Figures 1 and 2. This cannot be explained in terms of a rigid band model, where the major effect of alloying is to gradually fill (or empty) the electron states. The variation of W_gT^2 can be explained in two ways.

(*a*) The variation of W_gT^2 arises from a variation in W_ET^2, in turn due to a variation in the density of electrons and in the strength of the electron-phonon interaction; however, the density of states does not vary according to the rigid band model, but is increased by the addition of solutes of either valency (i.e. palladium and cadmium in silver, nickel, and zinc in copper both increase the density of electron states and thus W_E).

---

* Recently Dr. J. E. Zimmerman (Ford Motor Co., Detroit) has obtained further data on dilute copper-nickel, and Dr. H. M. Rosenberg (Oxford) on dilute copper-zinc.

(*b*) It is possible that the observed values of W_gT^2 contain an appreciable dislocation component W_DT^2 which varies with composition. Although the specimens have been annealed above the recrystallization temperature, it is possible that they are not sufficiently free of dislocations, but that the solute atoms lock and retain many dislocations, their concentration varying with solute content.

In spite of these variations, one can still estimate W_ET^2 for the parent metal by extrapolation, and the values thus obtained are confirmed by those obtained from alloys containing a small amount of iron (White, Woods, and Elford 1959). It is clear that the results favour the Makinson coupling scheme, i.e. the conduction electrons in copper, silver, and gold interact with transverse waves nearly as strongly as with longitudinal ones.

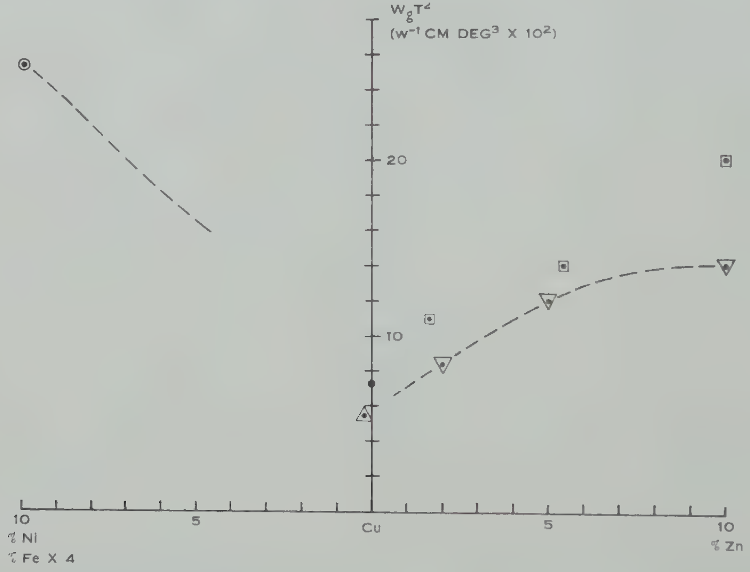


Fig. 2.—Values for W_gT^2 at liquid helium temperatures for various annealed copper alloys : ⊙, Cu-Ni (Estermann and Zimmerman 1952) ; △, Cu-Fe (White, Woods, and Elford 1959) ; ⊡, Cu-Zn annealed at 500 °C ; ▽, Cu-Zn annealed at 850 °C (Kemp, Klemens, and Tainsh 1957) ; ●, theoretical value for W_ET^2 (Klemens 1954*a*).

IV. Dislocations in Annealed Alloys

Since this variation cannot be explained as a variation of W_ET^2 in terms of the rigid band model, it has been assumed that it is due to an appreciable dislocation component W_DT^2. We have attempted to reduce this dislocation density by annealing the alloy specimens at higher temperatures, well above the recrystallization temperature.

In the case of three copper-zinc alloys (2, 5, and 10% Zn), an increase of the annealing temperature from 500 to 850 °C did indeed reduce W_gT^2 appreciably. If the difference between W_gT^2 and the theoretical value of W_ET^2 is entirely due to dislocations, then this increase in annealing temperature caused the dislocation density to be approximately halved. Similarly, in the case of Au-1·8% Pt, an increase in annealing temperature from 700 to 900 °C caused a very small change in W_gT^2 only, but a further increase to 1080 °C reduced the dislocation density by a factor of about 2/3.

Nevertheless, even these lowered values of W_gT^2 still show a variation with alloy composition. This indicates either that the residual dislocation densities are still controlled by a solute-dependent locking mechanism or that there is a variation of W_ET^2 with composition superimposed on the dislocation effect.

Furthermore, there are quite a number of alloys where the removal of dislocations is very sluggish (e.g. Au-6·5% Cr) or where the effect of increasing the annealing temperature is very small (e.g. Cu-1% Si, Cu-1% Pt, Au-0·65% Cr, Au-0·7% Pt). These results certainly throw some doubt on the dislocation mechanism as the sole cause of the variation of W_gT^2 with composition.

In Table 1 we have summarized the values of W_gT^2 for various alloys, the effect of increasing the annealing temperature, and the dislocation densities calculated from the theory of phonon scattering by dislocations, ascribing the entire deviation of W_gT^2 from the pure metal value to W_DT^2. The dislocation resistivity is given by

$$W_DT^2=0\cdot17(vh^2/K^3)Nb^2 \qquad \text{(1)}$$

$$\left.\begin{aligned}&=5\cdot1\times10^{-9}N, \text{ for copper,}\\&=2\cdot0\times10^{-9}N, \text{ for silver,}\\&=4\cdot0\times10^{-9}N, \text{ for gold,}\end{aligned}\right\} \qquad \text{(2)}$$

where v is the phonon velocity, N the dislocation density (in lines/cm²), B the Burgers vector, h and K the Planck and Boltzmann constants and in (2) W_D is expressed in W^{-1} cm deg.

We shall see in Section VI that the theory overestimates N by a factor of between 3 and 6. The densities of Table 1 have thus been appropriately reduced. Furthermore, one cannot at present be certain of the terminal values of W_DT^2 and the corresponding values of N; though changes in W_gT^2 must be due to dislocations. Nevertheless, it seems clear that alloys of moderate solute content commonly contain, even after annealing above the recrystallization temperature, dislocation densities of the order of 10^{10} to 10^{11} lines/cm². These densities are not very much lower than the densities induced in the same alloys by severe plastic deformation.

This result seems, at first sight, somewhat surprising, but it should be remembered that solute atoms can retain dislocations in two ways. They can lock them, in the sense that considerable stresses are required to make them mobile, and they can reduce their average velocity under stress by many orders of magnitude. This has been demonstrated in the case of lithium fluoride (Johnston and Gilman 1959).

V. VARIATION OF DENSITY OF ELECTRONS

Although the observed variations of W_gT^2 are not consistent with a rigid band model, and although there is good evidence that W_D is frequently not negligible, one cannot rule out the possibility that the variation of W_gT^2 is at least partly due to a change of W_ET^2 due to a change of band structure on alloying. One would expect some correlation between changes in the electron density and in $W_E T^2$. In the case of copper and its alloys it appears from magnetic susceptibility measurements (Pugh and Ryan 1958) and electronic specific heat observations (Rayne 1957 ; Guthrie, Friedberg, and Goldman 1959) that the density of states increases with the addition both of nickel and of zinc.

TABLE 1

LATTICE THERMAL RESISTIVITIES AND DISLOCATION DENSITIES

Alloy	Ann. Temp. (°C)	$W_g T^2$	$W_D T^2$	Difference*	N	Difference*
		(10^2 W^{-1} cm deg^3)			(10^{10} cm^{-2})	
Cu (theor. value for W_E)		7·0				
Cu-2% Zn	500	11	4	2·7	1·3	0·8
	850	8·3	1·3		0·5	
Cu-10% Zn	500	20	13	6	4·2	2·0
	850	14	7		2·2	
Cu-7·5% Au	750	18	11		3·6	
Cu-2·3% Si	700	25	18	4	6·0	1·5
	1000	21	14		4·5	
Cu-1% Pt	700	15·4	8	1	2·6	0·3
	1030	14·5	7		2·3	
Cu-0·4% As	450	15	8		2·6	
Ag (theor. value for W_E)		4·3				
Ag-2% Pd	610	7·8	3 (?)		2·5 (?)	
Ag-5% Pd	610	10·0	5 (?)		4 (?)	
Ag-2% Cd	530	5·2	1		0·8	
Ag-5% Cd	530	6·7	2·5		2	
Au (theor. value for W_E)		3·5				
Au-0·7% Pt	750	3·9	0·4 (?)	nil	0·2 (?)	nil
	1050	3·9	0·4 (?)		0·2 (?)	
Au-1·8% Pt	700	6·6	3·1	0·2	1·3	0·1
	900	6·4	2·9	0·8	1·2	0·3
	1080	5·6	2·1		0·9	
Au-0·65% Cr	500	9·1	5·6	0·8	2·3	0·3
	1050	8·3	4·8		2·0	
Au-6·2% Cr	200	71	65 (?)	25	27	10
	500	46	40 (?)		17	

* Differences refer to changes on further annealing, where applicable.

Guthrie (1959) has recently found an increase in the electronic specific heat of copper with equal amounts of nickel and zinc added, so as to keep the electron concentration constant. The corresponding measurement of $\varkappa_g$ has not yet been done, but Kemp, Klemens, and Tainsh (1957) measured $\varkappa_g$ of two copper-gold alloys and found an increase in $W_g T^2$.

It is hoped that measurements of the low temperature thermoelectric power will provide a means of distinguishing between the effects of dislocations and electrons on the heat transport by phonons. In addition to the normal electronic thermoelectric power, which in an alloy should vary as T, since $\rho_0 \gg \rho_i$, there should be a component due to " phonon drag " S_g, proportional to T^3 at very low temperatures (Klemens 1954*b*). The magnitude, in the case of scattering of phonons by electrons and by dislocations only, is given by

$$S_g = -(C_g/3Ne)\varkappa_g W_E, \qquad \ldots\ldots\ldots\ldots\ldots\ (3)$$

where N is the number of electrons per unit volume, $C_g \propto T^3$ is the lattice specific heat per unit volume, and $-e$ the electronic charge. It should thus be possible to deduce separately the magnitudes of W_E and W_D from simultaneous observations of $\varkappa_g$ and S_g.

VI. Dislocation Densities in Deformed Alloys

While equations (**1**) and (**2**) allow the densities of dislocations to be estimated from the lattice thermal conductivities, these estimates must be treated with reserve, since the theoretical scattering cross sections, on which (**1**) is based, depend upon various approximations, of which the most vulnerable is probably the estimation of the anharmonicity coefficients according to Grüneisen's theory. It had been suspected for some time that the theory underestimates the scattering cross sections and therefore overestimates the dislocation densities (e.g. Klemens 1958). It is thus desirable to measure W_D on specimens for which an independent estimate of dislocation densities is available.

While Lomer and Rosenberg (1959) compared lattice thermal conductivity measurements on deformed Cu-30% Zn with dislocation counts on an electron micrograph, we have compared changes in the lattice thermal conductivity with stored energy release on annealing.

Measurements were made on Cu-0·4% As after severe torsional deformation and after annealing up to 175, 275, and 450 °C, the annealing temperature increasing uniformly at a rate of 6°/min. The same material had been deformed and annealed in the same manner by Clarebrough, Hargreaves, and West (1955), who studied the stored energy release on annealing. The only change in $\varkappa_g$ noted was a decrease in W_D on annealing between 275 and 450 °C; according to (**1**) or (**2**) this corresponded to the removal of 9×10^{11} dislocations/cm². From the stored energy release, correcting for the fact that the long-range energy of the dislocations was evidently released below 275 °C, one would estimate the density of dislocations removed as between 2 and 3×10^{11} cm^{-2}. Thus for copper (**1**) seems to over-estimate the dislocation density by a factor between 3 and 4·5 (Kemp, Klemens, and Tainsh 1959).

A similar study was made on deformed Cu-30% Zn, the stored energy of which was measured similarly by Clarebrough, Hargreaves, and Loretto (1960) both after quenching and after deformation. Here the dislocation density removed on annealing between 250 and 400 °C from changes in W_gT^2 appeared, from (**1**), to be about $2\cdot6\times10^{12}$ cm^{-2}; the stored energy (without correction for a possible early release of the long-range strain energy) indicated 4×10^{11} cm^{-2}; the discrepancy is thus about 6·5, but may be less because of uncertainty in the stored energy data.

The case of Cu-30% Zn is complicated by a large release of energy below 250 °C which, according to Clarebrough, Hargreaves, and Loretto (1960), is due to the removal of point defects (probably of the order of 1–2 per cent. vacancies), and the change in the lattice conductivity at liquid oxygen temperatures indicates the removal of vacancies of comparable concentration.

In view of these results, and those of Lomer and Rosenberg (1959), it seems fairly certain that estimates of dislocation densities from the lattice thermal conductivity using equation (**1**) are too high by a factor of between 3 and 6. Correspondingly one can now estimate dislocation densities in face-centred cubic metals from their lattice thermal conductivity to within an uncertainty factor of 2.

VII. REFERENCES

BIRCH, J. A., KEMP, W. R. G., KLEMENS, P. G., and TAINSH, R. J. (1959).—*Aust. J. Phys.* **12**: 455.

CHARI, M. S. R., and DE NOBEL, J. (1959).—*Physica* **25**: 60, 84.

CLAREBROUGH, L. M., HARGREAVES, M. E., and LORETTO, M. (1960).—*Proc. Roy. Soc.* A. (In press.)

CLAREBROUGH, L. M., HARGREAVES, M. E., and WEST, G. W. (1955).—*Proc. Roy. Soc.* A **232**: 252.

COHEN, M. H., and HEINE, V. (1958).—*Advanc. Phys.* **7**: 395.

ESTERMANN, I., and ZIMMERMAN, J. E. (1952).—*J. Appl. Phys.* **23**: 578.

GUTHRIE, G. L. (1959).—*Phys. Rev.* **113**: 793.

GUTHRIE, G. L., FRIEDBERG, S. A., and GOLDMAN, J. E. (1959).—*Phys. Rev.* **113**: 45.

JOHNSTON, W. G., and GILMAN, J. J. (1959).—*J. Appl. Phys.* **30**: 129.

KEMP, W. R. G., KLEMENS, P. G., SREEDHAR, A. K., and WHITE, G. K. (1956).—*Proc. Roy. Soc.* A **233**: 480.

KEMP, W. R. G., KLEMENS, P. G., and TAINSH, R. J. (1957).—*Aust. J. Phys.* **10**: 454.

KEMP, W. R. G., KLEMENS, P. G., and TAINSH, R. J. (1959).—*Phil. Mag.* **4**: 845.

KLEMENS, P. G. (1954*a*).—*Aust. J. Phys.* **7**: 57.

KLEMENS, P. G. (1954*b*).—*Aust. J. Phys.* **7**: 520.

KLEMENS, P. G. (1955).—*Proc. Phys. Soc. Lond.* A **68**: 1113.

KLEMENS, P. G. (1956).—" Encyclopedia of Physics." Vol. 14, Pt. 4. (Springer: Berlin.)

KLEMENS, P. G. (1958).—*Solid State Physics* **7**: 1.

KLEMENS, P. G. (1959).—*Aust. J. Phys.* **12**: 199.

LOMER, J. N., and ROSENBERG, H. M. (1959).—*Phil. Mag.* **4**: 467.

MAKINSON, R. E. B. (1938).—*Proc. Camb. Phil. Soc.* **34**: 474.

PUGH, E. W., and RYAN, F. M. (1958).—*Phys. Rev.* **111**: 1038.

RAYNE, J. A. (1957).—*Phys. Rev.* **108**: 22.

SOMMERFELD, A., and BETHE, H. (1933).—" Handbuch der Physik." Vol. 34/2, Pt. 3. (Springer: Berlin.)

WHITE, G. K. (1960).—*Aust. J. Phys.* **13**: 255.

WHITE, G. K., WOODS, S. B., and ELFORD, M. T. (1959).—*Phil. Mag.* **4**: 688.

THERMAL TRANSPORT IN DILUTE ALLOYS

By G. K. White*

[*Manuscript received September* 1, 1959]

Summary

Experimental values of the thermal resistivity for a number of elements, including copper, silver, gold, argon, germanium, silicon, antimony, and bismuth, are compared with theoretical values deduced from the anharmonic coupling between lattice vibrations. For copper-, silver-, and gold-rich alloys experimental and theoretical data for scattering of lattice waves by point defects are compared.

I. Introduction

Whereas some other papers on heat conduction at this conference (see, for example, Klemens (1960) and Kemp and Klemens (1960)) discuss the role played by free electrons and by dislocations in scattering lattice waves at low temperatures, this paper considers some of the experimental evidence concerning the mutual interaction of phonons and interaction of phonons with point defects, which is available from the thermal conductivity measurements of the author and his colleagues.

At temperatures near and above $\theta_D/5$, θ_D being the Debye temperature, theoretical considerations (see, for example, review of Klemens 1958) indicate that the two scattering processes due to anharmonic interactions between lattice waves and to scattering of lattice waves by point defects are the major factors in limiting the lattice thermal conductivity λ_g, each process being responsible for a thermal resistivity (W_U and W_P respectively) which varies directly as temperature. Thus

$$1/\lambda_g = W_g \simeq W_U + W_P \propto T,$$

whence

$$\lambda_g \propto T^{-1}.$$

Indeed, many observations between 50 and 100 °K on metals, semiconductors, and dielectric crystals have confirmed that this temperature dependence is approximately correct.

II. Anharmonic Interaction

It is of interest to compare the magnitudes of the resistivity observed experimentally with those deduced from approximate theoretical treatments. Studies on pure semiconductors, in which the electronic conductivity λ_e may be negligible, and on pure dielectric crystals give values of λ_g and therefore of W_U directly. However, in pure metallic elements the lattice conductivity is masked by the presence of the electronic component, and it is only in suitable alloys that values of λ_g can be deduced from experiment. Such alloys must contain solute atoms which can scatter electrons strongly and thereby reduce λ_e, but, if values of W_U are required, it is necessary that they should not scatter the lattice waves too strongly (so that $W_P \ll W_U$). In the solvents copper, silver, and gold the heterovalent solutes, such as iron, arsenic, germanium, tin, present in the proportion of about one atom per thousand fulfil these conditions fairly well.

---

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

For specimens of copper (with 0·06 atomic per cent. iron), silver (+0·1 and 0·3 at. % Sn) and gold (+0·1 at. % Fe) it is observed (White and Woods 1954; White, Woods, and Elford 1959) that λ_g has magnitudes of $35/T$, $14/T$, $8/T$ W cm^{-1} deg^{-1} respectively, and it is believed that the reciprocals of these magnitudes represent reliable values for the anharmonic interaction term W_U. Data obtained in Sydney on rather less dilute alloys of silver (containing cadmium or palladium), gold (+platinum or chromium), and copper (+zinc) when extrapolated to zero percentages of solute seem to confirm these magnitudes.

TABLE 1

COMPARISON OF LATTICE THERMAL CONDUCTIVITIES LIMITED BY ANHARMONIC COUPLING

Element	θ_D (°K)	$\lambda\gamma^2T$ (Theoretical) (W cm^{-1})	λT (Experimental) (W cm^{-1})	γ_{EXP}	γ_G
Cu	310	246	35	2·6	2·0
Ag	220	169	14	3·4	2·4
Au	185	143	8	4·2	3·0
Ne	64	0·8	0·075	3·3	2·9 (?)
A	80	3·9	0·25	3·9	2·9
Kr	63	4·3	0·27	4·0	2·9 (?)
Sb	210	183	34	2·3	0·9
Bi	120	62	14	2·1	1·1
Si	700	1880	500	1·9	0·7
Ge	400	948	200	2·2	0·7
Te	180	121	10	3·5	0·8
Se	250	191	~6	~5	—

Theoretically, the magnitude of the anharmonic coupling and hence of W_U (or λ_U) has been deduced for a simplified model of a cubic crystal. By generalizing the analysis for a linear chain and assuming central forces and nearest neighbour interaction, Leibfried and Schlömann (1954) deduced that

$$\lambda \simeq 3 \cdot 61 aA\theta_D^3/\gamma^2 T \quad \text{W cm}^{-1}\,\text{deg}^{-1}, \qquad \ldots\ldots\ldots\ldots \quad (1)$$

where a is the lattice constant, A is the atomic weight, θ_D is the Debye temperature, and γ is the Gruneisen parameter. Essentially equivalent expressions have been given by Dugdale and MacDonald (1955) and by Kontorova (1957). In Table 1 are given values of $\lambda\gamma^2T$ deduced from this expression, together with experimental values of λT. (Latter values are from White and Woods 1954, 1956, 1958; Fischer, White, and Woods 1957; White, Woods, and Elford 1958, 1959.) A difficulty in comparing experiment and theory lies in choice of θ_D and of γ, each of which affects λT strongly. For the computations of Table 1, values of θ_D deduced from specific heat data at normal temperatures are used. γ may be calculated from the expansion coefficient, specific heat, and compressibility, and those values marked in the table as γ_G are calculated in this way (see, for example, the tabulation of Gruneisen (1926)). However, it is questionable, particularly for anisotropic materials, whether this represents a suitably averaged γ for substitution in equation (**1**). Barron (1956) has pointed out that in anisotropic elements for which $\bar{\gamma}$ (calculated from expansion data) is small, the variations of γ among the individual normal modes cannot be neglected. Thus in the limit where $\bar{\gamma} \leqslant 0$, there will still be cubic terms in the crystal potential causing coupling of the modes and hence causing a thermal resistance.

The values of γ_{EXP} calculated from the ratio of $\lambda\gamma^2 T$ (theoretical) and λT (experimental) are shown in the table; comparison with the values γ_G indicates that a discrepancy of about 30–40 per cent. (corresponding to a factor of 2 in the conductivity) exists for the cubic elements, and for antimony, bismuth, silicon, etc., the discrepancy is much greater.

Table 2

VALUES OF W/T IN CM W^{-1}

Metal	% Solute		Observed W/T	W_U/T (Table 1)	W_P/T (Observed minus W_U/T)	W_P/T (eqn. (**2**), (**3**))
Cu	2	Zn	0·040	0·0285	0·012	6×10^{-5}
	5	Zn	0·071	0·0285	0·043	15×10^{-5}
	7·5	Au	0·40	0·0285	0·375	0·5
	0·5	Pt	$\simeq$0·13	0·0285	$0{\cdot}10_5\pm0{\cdot}02$	$0{\cdot}07_3$
	1·0	Pt	0·18	0·0285	0·15	0·145
Au	0·7	Pt	0·16	0·125	0·035	2×10^{-6}
	1·8	Pt	0·216	0·125	0·09	4×10^{-6}
	0·65	Cr	0·142	0·125	0·02	0·0064
	6·2	Cr	0·25	0·125	0·125	0·06

III. Point Imperfections

Turning to the scattering of lattice waves by point defects, which also results in a thermal resistance varying directly as temperature, the theoretical relation (see Klemens 1955) leads to

$$W_P/T = 3{\cdot}3\times10^{11}(nS^2/v^2)\cdot(A/\rho) \quad \text{cm W}^{-1}, \qquad \ldots\ldots\ldots \ (\mathbf{2})$$

where n is the relative defect concentration and ρ is the density. For scattering due to mass difference only (average mass M, and mass difference ΔM)

$$S^2 = \frac{1}{12}\left(\frac{\Delta M}{M}\right)^2, \text{ if } n\Delta M \ll M, \qquad \ldots\ldots\ldots\ldots \ (\mathbf{3})$$

or more generally

$$nS^2 = \frac{1}{12}\sum_i n_i(M_i - M)^2/M^2.$$

Now for copper, silver, and gold respectively containing 1 at. % of a solute of mass $M+\Delta M$, W_P/T may be calculated from (**2**) and (**3**) to have values of

$0\cdot034(\Delta M/M)^2$, $0\cdot056(\Delta M/M)^2$, and $0\cdot018(\Delta M/M)^2$. In Table 2 these theoretical values are compared with some experimental data (see Kemp *et al.* 1957 and Birch *et al.* 1959, and more recent unpublished work).

It may be noted that experimental data on binary systems in which the solute atom is a neighbour or near neighbour of the solvent atom, such as copper-zinc, gold-platinum, and the silver-cadmium and silver-palladium alloys, all yield thermal resistances much larger than might be expected if only mass defect scattering were to occur. Recently we have investigated the case of copper-platinum (Fig. 1), in which the mass difference ΔM is large and in which mass defect scattering might be expected to predominate, and here experiment and theory agree rather well. Fair agreement was obtained with the Cu+7·5% Au (Birch *et al.* 1959) but in alloys of this type a rather large percentage of solute is required to reduce the electronic conductivity and the perturbation treatment of the mass defect scattering may no longer be strictly valid.

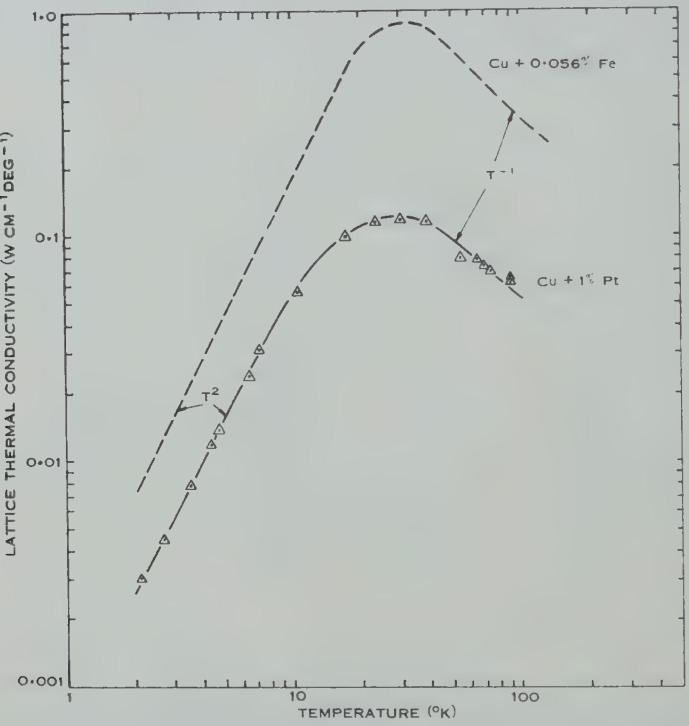


Fig. 1.—Lattice thermal conductivity λ_g of a copper+1 atomic per cent. platinum alloy, together with λ_g for " pure " copper (containing only 0·056% Fe, from White and Woods (1954)).

As Klemens (1955) has pointed out, scattering of the lattice vibrations may arise not only from relative mass changes at the defect site but from a general distortion or strain field surrounding the defect. In more general form than appears in equation (3), his theory shows that

$$S^2=S_1^2+(S_2+S_3)^2,$$

where

$$S_1=\Delta M/M\sqrt{12},$$

and

$$S_2=\sqrt{(2/3)}.\Delta v/v$$
$$S_3=-6\sqrt{(2/3)}.\Delta R/R.$$

Here $2\Delta v/v$ is the fractional change of the elastic constants of the nearest linkages and $\Delta R/R$ is the relative (outward) displacement of the neighbours nearest to the defect. An experimental test of this "distortion" scattering is rather difficult, as the terms S_2 and S_3 are much more uncertain than is S_1. It is apparent that in the alloys copper-zinc, gold-platinum these terms must play a major role in determining the resistance.

It is hoped to investigate such systems as potassium + sodium, lithium + magnesium, silver + platinum, silver + lead, gold + aluminium, aluminium + lead, and bismuth-rich alloys, as well as the solid binary mixtures of inert gases which we are examining at present.

IV. ACKNOWLEDGMENTS

The author is most grateful to Mr. R. J. Tainsh for his help with experiments and to Dr. P. G. Klemens for many useful discussions.

V. REFERENCES

BARRON, T. H. K. (1956).—*Nature* **178**: 871.

BIRCH. J. A., KEMP. W. R. G., KLEMENS, P. G., and TAINSH, R. J. (1959).—*Aust. J. Phys.* **12**: 455.

DUGDALE, J. S., and MACDONALD, D. K. C. (1955).—*Phys. Rev.* **98**: 1751.

FISCHER, G., WHITE, G. K., and WOODS, S. B. (1957).—*Phys. Rev.* **106**: 480.

GRUNEISEN, E. (1926).—"Handbuch der Physik." Vol. 10. p. 1. (Springer: Berlin.)

KEMP, W. R. G., and KLEMENS, P. G. (1960).—*Aust. J. Phys.* **13**: 247.

KEMP, W. R. G., KLEMENS, P. G., TAINSH, R. J., and WHITE, G. K. (1957).—*Acta Met.* **5**: 303.

KLEMENS, P. G. (1955).—*Proc. Phys. Soc. Lond.* A **68**: 1113.

KLEMENS, P. G. (1958).—*Solid State Physics* **7**: 1.

KLEMENS, P. G. (1960).—*Aust. J. Phys.* **13**: 238.

KONTOROVA, T. A. (1957).—*Soviet Physics, Technical Phys.* (Trans.) **1**: 1959.

LEIBFRIED, G., and SCHLÖMANN, F. (1954).—*Nachr. Akad. Wiss. Gottingen* **11a**: 71.

WHITE, G. K., and WOODS, S. B. (1954).—*Phil. Mag.* **45**: 1343.

WHITE, G. K., and WOODS, S. B. (1956).—*Phys. Rev.* **103**: 569.

WHITE, G. K., and WOODS, S. B. (1958).—*Phil. Mag.* **3**: 342, 785.

WHITE, G. K., WOODS, S. B., and ELFORD, M. T. (1958).—*Phys. Rev.* **112**: 111.

WHITE, G. K., WOODS, S. B., and ELFORD, M. T. (1959).—*Phil. Mag.* **4**: 688.

THE INTENSITY OF DIELECTRIC ABSORPTION IN ALKALI HALIDES AS A FUNCTION OF THE CONCENTRATION OF DIVALENT CATION IMPURITIES

By J. S. Cook* and J. S. Dryden*

[*Manuscript received September* 25, 1959]

Summary

The dielectric absorption of sodium chloride containing calcium and potassium chloride containing strontium has been measured as a function of concentration of the divalent impurity. From the intensity of the absorption it is possible to estimate accurately the number of divalent impurity cations associated with positive ion vacancies so as to form dipoles.

In both systems at equilibrium at room temperature it is found that a large proportion of the divalent impurity ions are associated with the positive ion vacancies in a higher complex than a dipole. By quenching from different temperatures, an investigation of the equilibria between divalent impurities, vacancies, dipoles, and higher complexes has been made. The kinetics of the aggregation of the dipoles has been studied for the $NaCl(Ca^{++})$ system.

I. Introduction

When a divalent ion replaces a positive ion in the lattice of an alkali halide a positive ion vacancy is introduced. A dipole is formed if the cation impurity, possessing an extra positive charge, and the cation vacancy, possessing an effective negative charge, are associated on near lattice sites. The dielectric absorption resulting from the rotation of this dipole has been studied by Burstein, Davidson, and Sclar (1954) (potassium chloride), Haven (1955) and Jacobs (1955) (sodium chloride), and Dryden and Meakins (1957) (lithium, sodium, and potassium halides). From measurements of the frequency at which this absorption occurs the activation energy for the dipolar rotation has been determined. The activation energy is that required to move a positive ion vacancy from one lattice site to an adjacent one.

The other quantity that can be measured is the intensity of the absorption, which is directly proportional to the number of dipoles ; that is, to the number of impurity ion-vacancy pairs present. Information about the degree of association of these pairs has been obtained previously from d.c. conductivity measurements of impure alkali halide crystals. The association has only a second-order effect on the conductivity, and deductions made from these measurements have only been approximate and not always consistent. (See e.g. Lidiard (1957).)

Dryden and Meakins (1957) reported the existence of a second region of dielectric absorption when the impurity concentration exceeded a certain amount, but they were unable from the nature of their samples to determine this concentration.

---

* Division of Electrotechnology, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

Haven (1955) made the observation, for sodium chloride containing calcium at about 100 °C, that the sum of the calcium ions associated in dipoles (calculated from the intensity of the dielectric absorption) and the number of free vacancies (calculated from the d.c. conductivity) was only a fraction of the number of calcium ions present. The rest were assumed to be present in higher complexes or in a segregated phase.

The present work was undertaken to obtain more complete information about the intensity of the dielectric absorption as a function of impurity concentration in NaCl containing Ca^{++} and in KCl containing Sr^{++}.

II. Experimental

Crystals 3–4 cm in length and 1·5–5 cm² in area were grown by pulling from the melt, at about 0·75 cm/hr, in an atmosphere (about 10 cm Hg) of pure nitrogen. The starting materials were B.D.H. reagent grade sodium chloride and Merck reagent grade potassium chloride and were contained in a graphite crucible. The sodium chloride was purified by melting, freezing slowly, and discarding the core several times. The divalent impurities added to the melt were the purest commercially available.

Samples of thickness 0·1–0·2 cm were cleaved from the crystals and were measured in a three-electrode assembly which could be evacuated. In some cases aluminium electrodes were evaporated on to the samples, otherwise corrections were made for the air gap. The samples were quenched by removing them from an oven to the room atmosphere or sometimes by placing them on a block of metal at room temperature.

The bridge used for the measurements has been described by Thompson (1956). Errors in the dielectric loss factor (ε'') are estimated to be ± 5 per cent. or $\pm 0{\cdot}001$, whichever is the greater.

Analysis of the calcium in the sodium chloride crystals was done by radioactivation* and the strontium in potassium chloride by fluorescent X-ray spectroscopy.†

III. Results and Discussion

(a) *Sodium Chloride containing Calcium*

Dielectric measurements were made on four crystals grown from melts containing the following amounts of calcium : none added ; 10^{-4} mole fraction ; 5×10^{-4} mole fraction ; 10^{-3} mole fraction. The concentration of calcium in the four crystals is shown in Table 1, from which it may be seen that the ratio of the concentration of calcium in the crystal to that in the melt is high.

Preliminary measurements revealed that the intensity of the dielectric absorption measured at room temperature (20 °C) did not have a definite value but was dependent on thermal history. In Figure 1 the dielectric absorption is plotted for a sample containing $3\cdot6\times10^{-4}$ mole fraction (*a*) after standing for some time at room temperature and (*b*) after quenching from 130 °C. The curve drawn through the experimental points in Figure 1 (*b*) is a Debye curve

---

* The radiochemical work was performed by the Isotopes Section of the Australian Atomic Energy Commission, Lucas Heights, the irradiation of the samples being done at A.E.R.E., Harwell.

† The authors are grateful for the assistance of the Analytical Chemistry Section of the A.A.E.C., Lucas Heights, in performing this analysis.

$$\varepsilon''=0\cdot300\omega\tau/(1+\omega^2\tau^2), \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

TABLE I

SODIUM CHLORIDE (CALCIUM)

Concentration of Ca^{++} in the Melt (mole fraction)	Concentration of Ca^{++} in the Crystal (mole fraction) (Errors 95% Confidence Limits)	$\varepsilon''_{\text{max.}}$ (Calculated)	$\varepsilon''_{\text{max.}}$ (Measured)
None added	$1\cdot9\pm3\cdot4\times10^{-5}$	$0\cdot008\pm0\cdot014$	$0\cdot0080\pm0\cdot001$
10^{-4}	$6\cdot8\pm4\cdot5\times10^{-5}$	$0\cdot028\pm0\cdot018$	$0\cdot036_5\pm0\cdot002$
5×10^{-4}	$2\cdot2\pm0\cdot41\times10^{-4}$	$0\cdot090\pm0\cdot017$	$0\cdot088\pm0\cdot005$
10^{-3}	$3\cdot64\pm0\cdot74\times10^{-4}$	$0\cdot150\pm0\cdot030$	$0\cdot150\pm0\cdot007$

where ω is the angular frequency and τ a relaxation time. This indicates that immediately after quenching from 130 °C all the dipoles have the same relaxation time. However, after storage at room temperature, the intensity of the absorption decreases and a second absorption appears at a higher frequency. It is concluded from this observation that association of the dipoles into a higher complex occurs on standing at room temperature and some or all of the dipoles in the higher complex contribute to the polarization with a shorter relaxation

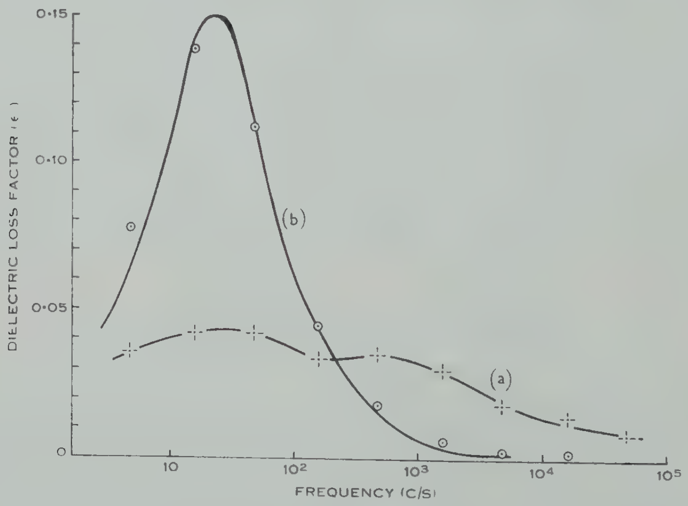


Fig. 1.—Dielectric absorption, measured at 20 °C, as a function of frequency for NaCl containing $3\cdot6\times10^{-4}$ mole fraction of Ca^{++} (*a*) after storage for some time at 20 °C and (*b*) immediately after quenching from 130 to 20 °C.

time. Schneider and Caffyn (1955) and Watkins (1959), using electron spin resonance, have observed a slow aggregation process involving Mn^{++} ions in NaCl and KCl.

The relation between the intensities in the two absorption regions is such that when the main absorption decreases by a certain amount m the second absorption increases by approximately $m/10$. On further storage the second peak increases in frequency and also in intensity until, as shown in Figure 1, the intensity approaches one-fifth of the original intensity of the main absorption.

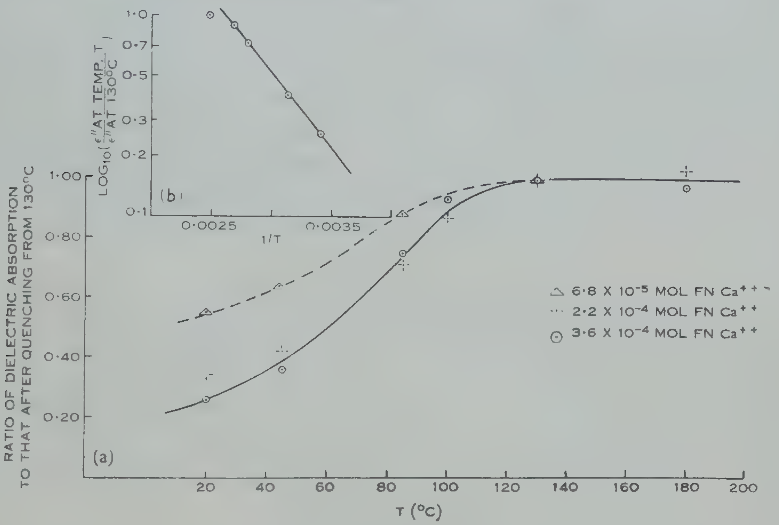


Fig. 2 (*a*).—Intensity of the dielectric absorption measured at 20 °C in NaCl(Ca) plotted as a function of the temperature from which the sample was quenched. The intensity of the absorption is expressed as a ratio to that after quenching from 130 °C.

Fig. 2 (*b*).—The full line is plotted on a logarithmic scale as a function of $1/T$.

(i) *Comparison with Simple Theory*.—The equilibria we are concerned with in a crystal of sodium chloride containing calcium ions may be represented by

$$\text{Ca}^{++}\text{ ion}+\text{Free Na}^{+}\text{ vacancy}\overset{\text{I}}{\rightleftharpoons}(\text{Ca}^{++}\text{-vacancy})\text{ Dipole}$$
$$\underset{\text{II}}{\rightleftharpoons}\text{Aggregates}+\text{Some free vacancies.}$$

If, as will be shown below, the rate at which equilibrium II is reached at room temperature is slow compared to I then by quenching from various temperatures it is possible to investigate equilibrium II. To do this samples were stored at different temperatures, quenched to room temperature, and the dielectric absorption measured within 10 min ; this was done with increasing temperature of storage. The results are shown in Figure 2 (*a*) plotted as a fraction of the dielectric absorption after quenching from 130° for each sample against the temperature of storage. The values shown for 20 °C for the two higher concentrations are approaching the equilibrium value but that for the lowest concentration is far from equilibrium.

The intensity of the absorption to be expected is given by the following expression

$$\varepsilon''_{\text{max.}}=p.2\pi e^2c/3akT, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{2})$$

where $\varepsilon''_{\text{max.}}$ is the maximum in the dielectric loss factor curve, c is the mole fraction of divalent ions, a is the inter-ion distance, and p is the fraction of the divalent ions present which are associated with a vacancy in the form of a dipole. The assumptions implicit in this expression are: (1) the dipole is formed by the vacancy being on the nearest neighbour site (it has been suggested on theoretical grounds that the next nearest site is more favourable energetically); (2) each ion carries a full electronic charge; and (3) there is no internal field correction. The $\varepsilon''_{\text{max.}}$ (calculated) values listed in Table 1 are calculated from equation (**2**) with $p=1$. The agreement between these values and the $\varepsilon''_{\text{max.}}$ measured after quenching from 130 °C or higher is good, from which it is concluded that at this temperature, within experimental error, none of the calcium ions are associated in complexes.

These measurements do not support theoretical calculations (Ninomiya 1959) designed to show why equation (**2**) should not apply.

The fraction of dipoles not in a higher complex at any particular temperature can be obtained from the full line of Figure 2 (*a*). This fraction appears to be independent of concentration.

Figure 2 (*b*) is a plot of $\log p$ as a function of $1/T$. It will be seen from the figure that $p=90 \exp(-E/kT)$ where $E=0{\cdot}15$ eV.

(ii) *Association Energy between an Impurity Ion and a Vacancy.*—If the divalent ion and the vacancy are regarded as associated only when they are nearest neighbours, then

$$p'/(1-p')^2=12c \exp(F_a/kT) \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{3})$$

for a NaCl lattice (Lidiard 1955), where F_a is the free energy of association and p' is the ratio of the Ca^{++} ions present in dipoles to the sum of those present in dipoles and as free Ca^{++} ions (i.e. on both sides of equilibrium I). In the derivation of this expression Coulomb attractions among the free defects have been ignored. An expression for $p'/(1-p')^2$ has been calculated allowing for this interaction, but for the values of F_a/kT involved in the NaCl(Ca) system the corrections to equation (**3**) are small.

It is assumed that at 130 °C and higher none of the Ca^{++} ions and vacancies are present in higher complexes, i.e. $p'=p$. Because of the comparatively large uncertainty in the chemical analysis, the highest value that can be assigned to p' is $0{\cdot}8$, but p' could be nearer unity. Substituting $p'=0{\cdot}8$ in equation (**3**) gives $F_a=10kT$. In the absence of experimental information about the rate at which equilibrium is established between unassociated defects and dipoles following quenching from 130 °C or higher there is uncertainty whether T can be taken as 130 °C or whether a lower value must be used; using 20 °C, F_a becomes

0·26 eV. The free energy of association is therefore 0·26 eV or higher. There have been several calculations, based on lattice theory, of the association energy, a recent one (Fumi and Tosi 1957) gives 0·38 eV for Ca^{++} in NaCl.

(iii) *Rate of Aggregation of Dipoles.*—When stored at room temperature after quenching from 130 °C the intensity of the dielectric absorption decays slowly with time at a rate dependent on concentration (Fig. 3). The observations have not been carried on for sufficient time to be certain that equilibrium has been reached for the most concentrated sample, but, because the measurement after 77 days falls on the straight line in Figure 2, it is believed that this sample is not far from equilibrium. Equilibrium obviously has not been reached in the two other samples.

The shape of the curves in Figure 3 approximates to those expected for a third-order reaction and, if so, the time taken to reach a given fraction of dipoles should be inversely proportional to the square of the initial concentration. The shift in the time scale of the three curves in Figure 3 is in the ratio of 1 : 3 : 18, whereas the inverse ratio of the squares of the concentrations is 1 : 2·6 : 17. In calculating this ratio the initial intensity of the dielectric absorption has been taken as a better measure of the initial concentration than the value from chemical analysis. Thus it appears that the kinetics of the aggregation of dipoles approximate to third-order in calcium ion concentration and would exclude as a mechanism of aggregation the diffusion of individual dipoles to nuclei such as dislocation lines.

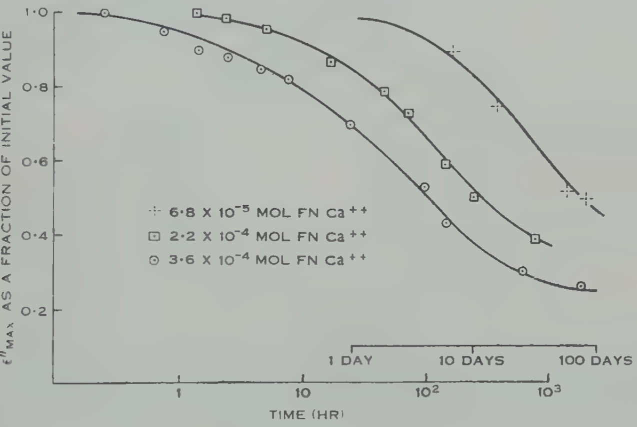


Fig. 3.—The decay of the intensity of the dielectric absorption in NaCl(Ca) as a function of the time following quenching from 130 to 20 °C

A sample from the crystal containing the highest concentration of Ca^{++} was stored at 42 °C after quenching from 130 °C, and the decay of dielectric absorption measured (Fig. 4). It can be seen that the time to reach 80 per cent. of the initial concentration is approximately one-half that required at 20 °C. Assuming that the activation energy for the diffusion of a dipole will be at least that for the rotation of a dipole, then the ratio of the diffusion rates of the dipole between these two temperatures would be at least 8 : 1. It would therefore appear that the rate of aggregation of the dipoles is not limited by the diffusion of the dipole, but, as the aggregation may proceed in two stages, as discussed below, this conclusion must be viewed with some caution.

The measurements at 42 °C have revealed one further piece of information which is not shown positively by the 20 °C measurements. The d.c. conductivity, as obtained from the intensity of the dielectric loss at frequencies below the absorption curve, increases as the decay in the intensity of the absorption curve proceeds. This indicates an increase in the concentration of free positive ion vacancies and means that in the aggregates there is not an exactly equal number of Ca^{++} ions and cation vacancies.

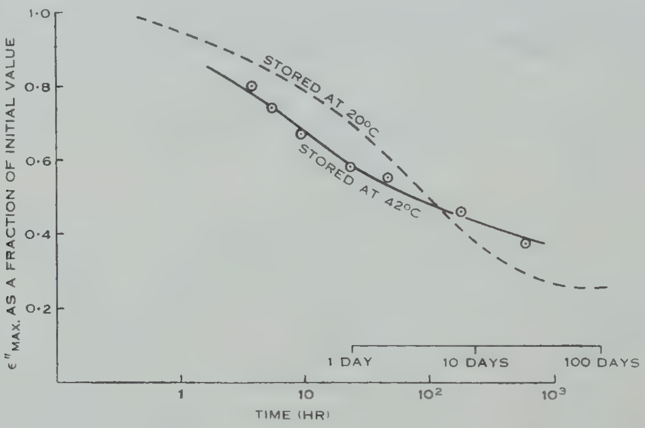


Fig. 4.—The decay of the intensity of the dielectric absorption in NaCl containing $3\cdot6\times10^{-4}$ mole fraction of Ca^{++} as a function of time following quenching from 130 to 42 °C. The measurements were made at 42 °C.

If the aggregation of dipoles consisted only of the union of three dipoles to form one unit then the dissociation of these aggregates would be described by the following equation :

$$\text{aggregates} \rightleftharpoons 3 \text{ dipoles}.$$

The equilibrium concentration of dipoles and aggregates at a given temperature would be described by the usual mass-action equation

$$\left.\begin{aligned} &[\text{dipoles}]^3/[\text{aggregates}]=\text{constant},\\ \text{or}\qquad &c^2p^3/(1-p)=\text{constant},\end{aligned}\right\} \qquad \ldots\ldots\ldots\ (4)$$

i.e. the fraction of dipoles at any temperature would be dependent on concentration. Because this is not the case (Fig. 2) then it is concluded that further aggregation into larger units must occur. This conclusion is supported by the changes which occur in the second absorption region, on storage, as mentioned earlier.

(*b*) *Potassium Chloride* (*Strontium*)

Crystals were grown from the melts shown in Table 2. An attempt to grow a crystal from a melt containing 10^{-2} mole fraction of Sr^{++} failed. The latter part of a crystal grown from the 5×10^{-3} mole fraction melt was completely cloudy with an inverted conical section of cloudy material extending upwards into the clear region. The two analyses shown for the solid grown from this latter concentration are (i) of a clear region of the crystal and (ii) of a region of the crystal containing a cloudy centre occupying 25 per cent. of the cross sectional area. The concentrations of Sr^{++} in the crystals grown from the lower concentrations are calculated from the measured $\varepsilon''_{\text{max.}}$ using equation (**2**).

TABLE 2

POTASSIUM CHLORIDE (STRONTIUM)

Concentration of Sr++ in the Melt (mole fraction)	Concentration of Sr++ in the Crystal (mole fraction)
10^{-5}	(4×10^{-6})
3×10^{-5}	$(1 \cdot 1 \times 10^{-5})$
10^{-4}	$3 \pm 1 \times 10^{-5}$
3×10^{-4}	$7 \pm 1 \cdot 5 \times 10^{-5}$
10^{-3}	$1 \cdot 4 \pm 0 \cdot 14 \times 10^{-4}$
3×10^{-3}	$4 \cdot 5 \pm 0 \cdot 2 \times 10^{-4}$
5×10^{-3}	(i) $9 \cdot 2 \pm 0 \cdot 5 \times 10^{-4}$
	(ii) $1 \cdot 4 \pm 0 \cdot 07 \times 10^{-3}$

It can be seen that potassium chloride crystals contain a considerably smaller fraction of the strontium in the melt than the equivalent fraction of calcium in sodium chloride. The potassium chloride used as starting material contained no detectable dielectric absorption, i.e. $\tan \delta \leqslant 0 \cdot 0001$ ($\varepsilon'' \leqslant 0 \cdot 0005$).

The dielectric absorption in these crystals is similar to that in sodium chloride (calcium), in that the intensity of the absorption measured at room temperature depends on the temperature from which the sample has been quenched. There are, however, some important differences.

Firstly, there is no trace of a second absorption region appearing just above the main absorption as aggregation of the dipoles occurs. All absorption curves measured are perfect Debye curves. If there is any mechanism of dielectric absorption in the aggregates it has a maximum at very different frequencies.

Secondly, the temperature at which the dissociation of the aggregates into dipoles is complete is concentration dependent. This has not been studied thoroughly but is illustrated by some measurements listed in Table 3. As discussed earlier this concentration dependence could indicate that the dipoles aggregate into small units and unlike Ca^{++} in NaCl no further aggregation occurs. No studies of the rate of decay of the dielectric absorption have been performed yet on these crystals but it is much slower at room temperature than for the same concentration of calcium in sodium chloride.

The final column in Table 3 is $\varepsilon''_{\text{max.}}$ calculated from equation (2) with $p=1$. It can be seen that, where the intensity of the absorption reaches a limit with increasing quenching temperature, agreement with the calculated value is good. This means that the dipoles consist of Sr^{++} ions and vacancies on nearest lattice sites. This is in contrast to the finding of Watkins (1959) from electron spin resonance studies that in the $KCl(Mn^{++})$ system next nearest lattice sites are favoured.

Because of the more accurate chemical analysis in this case as compared with sodium chloride, the lower limit which can be placed on the free energy of association using equation (3) is slightly higher at 0·28 eV.

Table 3

POTASSIUM CHLORIDE (STRONTIUM)

Concentration of Sr^{++} (mole fraction)	Intensity of Dielectric Absorption ($\varepsilon''_{\text{max.}}$)				$\varepsilon''_{\text{max.}}$ (Calculated)
	After Quenching from 80 °C	After Quenching from 130 °C	After Quenching from 185 °C	After Quenching from 350 °C	
3×10^{-5}	—	0·011	—	—	$0\cdot011\pm0\cdot003$
7×10^{-5}	—	—	0·021	—	$0\cdot026\pm0\cdot008$
$1\cdot4\times10^{-4}$	—	0·045	0·052	0·056	$0\cdot052\pm0\cdot005$
$4\cdot5\times10^{-4}$	0·016	0·022	0·069	0·095	$0\cdot169\pm0\cdot008$
$9\cdot2\times10^{-4}$	—	—	0·068	—	$0\cdot35\pm0\cdot018$
$1\cdot4\times10^{-3}$	—	0·024	0·074	—	$0\cdot52\pm0\cdot025$

IV. CONCLUSIONS

(1) In the two systems studied in equilibrium at room temperature a high percentage of divalent impurity and positive ion vacancies are present in higher complexes than pairs. With increasing temperature these complexes dissociate into pairs.

(2) When quenched from a temperature above which dissociation is complete, the intensity of the dielectric absorption agrees with simple theory.

(3) A lower limit has been obtained for the association energy between Ca^{++} and a positive ion vacancy in NaCl and that between Sr^{++} and a positive ion vacancy in KCl.

(4) There are significant differences in properties between the NaCl(Ca) and KCl(Sr) crystals. In NaCl(Ca) a second absorption region is detected but not in KCl(Sr). The aggregation of dipoles in KCl(Sr) appears to be simpler than the aggregation of dipoles in NaCl(Ca).

V. REFERENCES

Burstein, E., Davidson, J. W., and Sclar, N. (1954).—*Phys. Rev.* **96**: 819. (Abstract only.)

Dryden, J. S., and Meakins, R. J. (1957).—*Disc. Faraday Soc.* **23**: 39.

Fumi, F. G., and Tosi, M. P. (1957).—*Disc. Faraday Soc.* **23**: 92.

Haven, Y. (1955).—Concentration and association of lattice defects in NaCl. Report of Conference on Defects in Crystalline Solids, July 1954, p. 261. (Phys. Soc.: London.)

Jacobs, G. (1955).—*Naturwissenschaften* **42** : 575.

Lidiard, A. B. (1955).—Dielectric relaxation in impure polar crystals. Report of Conference on Defects in Crystalline Solids, July 1954, p. 283. (Phys. Soc. : London.)

Lidiard, A. B. (1957).—" Handbuch der Physik." (Ed. S. Flugge.) Vol. 20, p. 246. (Springer-Verlag : Berlin.)

Ninomiya, T. (1959).—*J. Phys. Soc. Japan* **14** : 30.

Schneider, E. E., and Caffyn, J. E. (1955).—Some observations of the paramagnetic resonance absorption, conductivity and optical absorption of single crystals of sodium chloride containing manganese. Report of Conference on Defects in Crystalline Solids, July 1954, p. 74. (Phys. Soc. : London.)

Thompson, A. M. (1956).—*Proc. Instn. Elect. Engrs.* **103B** : 704.

Watkins, G. D. (1959).—*Phys. Rev.* **113** : 79.

HIGH TEMPERATURE DIELECTRIC BREAKDOWN OF ALKALI HALIDES

By J. J. O'Dwyer*

[*Manuscript received September 25, 1959*]

Summary

In the high temperature region much of the experimental work on the dielectric breakdown of the alkali halides is apparently conflicting. If, however, it is assumed that breakdown is thermal in nature instead of intrinsic, reasons can be given which reduce the difference between various sets of existing experimental results. A calculation of the thermal breakdown strength is given based on the assumption that the electrical conductivity is principally ionic. The magnitude and temperature variation of the breakdown strength is given correctly without disposable constants. Some suggestions are given for experimental work which may clear up outstanding difficulties.

I. Introduction

The dielectric breakdown strength of single crystals of alkali halides a fraction of a millimetre in thickness has been investigated by various authors (cf. Austen and Whitehead 1940; Alger and von Hippel 1949; Calderwood and Cooper 1953; and Konorova and Sorokina 1957). Generally speaking, the breakdown strength increases with temperature up to about room temperature and then decreases with further increase in temperature. Breakdown is, moreover, commonly believed to be intrinsic in nature, i.e. an electronic instability initiates the breakdown process.

This relatively simple situation is disturbed by allegedly conflicting experimental facts in the high temperature region. Thus von Hippel and Alger (1949) and Konorova and Sorokina (1957) find that d.c. breakdown tests give results differing from those for impulse tests and that in addition the breakdown strength depends on the time of the voltage impulse (cf. Fig. 1). On the other hand Calderwood and Cooper (1953) find no significant difference between d.c. and impulse tests (cf. Fig. 2).

It is the purpose of this paper to propose a mechanism of thermal breakdown for alkali halides in the high temperature region, which mechanism, besides giving correct orders of magnitude to the breakdown strength and its temperature dependence, makes some reconciliation in the apparent experimental conflict.

---

* Department of Physics, The University of New South Wales, Sydney.

II. The Thermal Breakdown Strength

The general equation governing thermal conditions in a dielectric is

$$C_v \frac{\mathrm{d}T}{\mathrm{d}t} - \mathrm{div}\,(\varkappa \text{ grad } T) = \sigma F^2, \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

where C_v is the specific heat per unit volume, $\varkappa$ and σ are the thermal and electrical conductivities, F the field strength, $\mathrm{d}T/\mathrm{d}t$ is the time derivative, and grad T is the space gradient of the temperature. Assuming that the dielectric is destroyed if the temperature of its hottest part exceeds T_m (not necessarily as high as the melting point), then the thermal breakdown occurs for a field strength and time of application for which this temperature is just exceeded. Since C_v, $\varkappa$, and σ are all functions of T, and σ is in addition an explicit function of F, the equation cannot be solved in general terms, but two limiting cases arise for which approximate solutions can be determined.

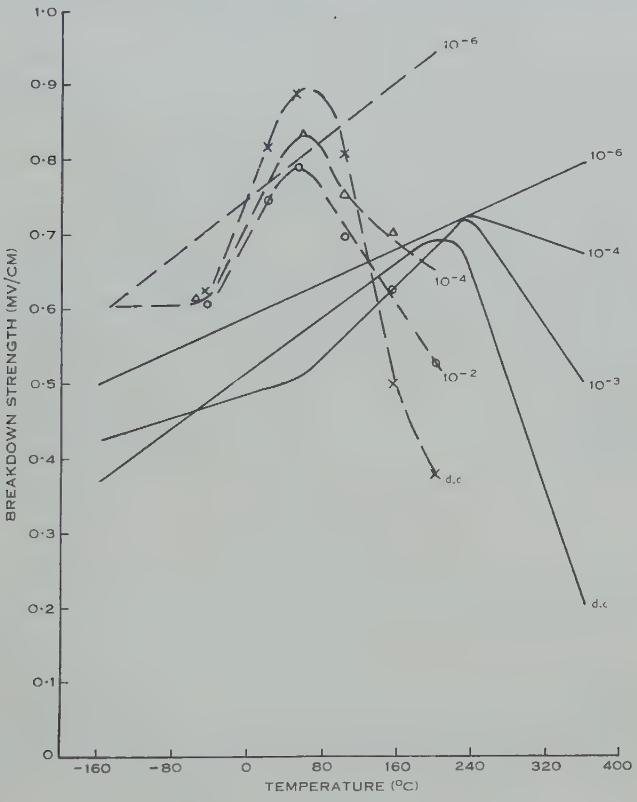


Fig. 1.—The dielectric breakdown strength of potassium bromide as a function of temperature for various rise times.
(——— Alger and von Hippel).
(------- Konorova and Sorokina).

The first approximation concerns the case in which the applied field is increased slowly and the heat conducted away sufficiently well so that the time-dependent term in (**1**) is always small. In these circumstances Whitehead (1950) has shown that a maximum thermal *voltage* V^* exists for a plane parallel slab of dielectric with electrodes on the opposite surface which are maintained at ambient temperature T_0. Whitehead's result is

$$V^{*2}=8\int_{T_0}^{T_m}\frac{\varkappa}{\sigma}\mathrm{d}T, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{2})$$

which can be evaluated for the alkali halides under certain simplifying assumptions. If we consider the conductivity to be mainly ionic, then we can use the result of Mott and Gurney (1948) for the conductivity due to the motion of lattice defects in an external field.† In the usual experimental result

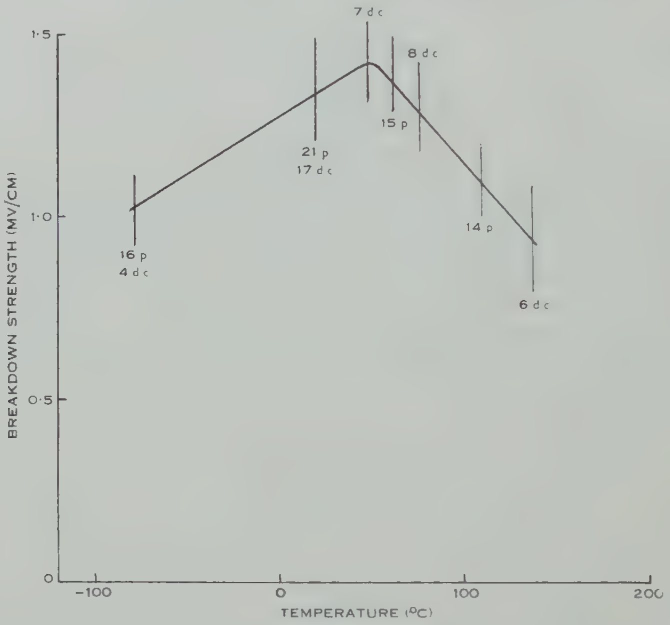


Fig. 2.—The dielectric breakdown strength of sodium chloride as a function of temperature for d.c. and impulse tests. The figures refer to the number of specimens tested d.c. or impulse (p). (Calderwood and Cooper).

$$\sigma=\sigma_0\exp\left(-\varphi/kT\right), \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{3})$$

with φ the activation energy of the process as measured, they found

$$\sigma_0=ANe^2\nu a^2/kT, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{4})$$

where A is a numerical factor, N the number of ion pairs per unit volume, e the electronic charge, ν the vibrational frequency, a the lattice constant, and k

---

† A theory of breakdown of alkali halides along these lines has been given by Franz (1952, 1956) considering only electronic conductivity. A similar theory has been given by Vermeer (1956) for glass.

Boltzmann's constant. Furthermore the thermal conductivity as a function of temperature is given by (cf. Klemens 1958)

$$\varkappa=\varkappa_0/T. \qquad (5)$$

Using (3), (4), and (5) in (2) we have

$$V^{*2}=8\int_{T_0}^{T_m}\frac{\varkappa_0 k}{ANe^2\nu a^2}\exp(\varphi/kT)\mathrm{d}T$$
$$\simeq 8\frac{\varkappa_0 k^2 T_0{}^2}{ANe^2\nu a^2\varphi}\exp(\varphi/kT_0), \qquad (6)$$

if $\varphi \gg kT$ for all T considered and $T_m > T_0$. Reintroducing σ_0 we have

$$V^*\simeq\left(\frac{8\varkappa_0 kT_0}{\sigma_0\varphi}\right)^{\frac{1}{2}}\exp(\varphi/2kT_0). \qquad (6a)$$

As will be shown later this voltage corresponds to a field strength far in excess of breakdown for the sample thicknesses normally used in laboratory experiments.

The second approximation is for the case in which the electrode system and the ambient medium are such poor thermal sinks that all of the heat generated goes to raise the temperature of the crystal. In these circumstances a maximum *field strength* exists which is dependent on the time variation of the applied field. We shall consider the relatively simple case in which the applied field is linear in the time, say

$$F=(F^*/t^*)t, \qquad (7)$$

where (anticipating the nature of the result) F^* is the breakdown field strength and t^* is the time taken to raise the field strength to breakdown. The equation for thermal breakdown is in this case

$$C_v(\mathrm{d}T/\mathrm{d}t)=\sigma F^2, \qquad (8)$$

and using (3), (4), and (7) in (8) we have

$$\frac{\mathrm{d}T}{\mathrm{d}F}=\frac{t^*}{F^*}\,\frac{ANe^2\nu a^2F^2}{C_v kT}\exp(-\varphi/kT). \qquad (9)$$

If $\varphi \gg kT$ the principal temperature dependence of (9) is in the exponential giving the approximation

$$\int_{T_0}^{T_m}\exp(\varphi/kT)\mathrm{d}T\simeq\frac{t^*}{F^*}\,\frac{ANe^2\nu a^2}{C_v kT_0}\int_0^{F^*}F^2\mathrm{d}F.$$

The integration with regard to T is again performed approximately as in (6) yielding

$$F^*=\left\{\frac{3C_v kT_0}{\sigma_0\varphi}\,\frac{1}{t^*}\right\}^{\frac{1}{3}}\exp(\varphi/2kT_0). \qquad (10)$$

It remains to estimate the order of the rise time of the voltage which separates " steady state thermal breakdown " of (**6a**) from " impulse thermal breakdown " of (**10**). This is easily obtained from (**1**) as

$$t^* \sim C_v l^2/\varkappa, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad \textbf{(11)}$$

where l is a length of order of the thickness of the dielectric specimen if the electrodes are good thermal sinks, but greater otherwise. For rise times less than that given by (**11**) equation (**10**) should hold, while for longer rise times (**6a**) is appropriate. For rise times of order of that given by (**11**) all terms in (**1**) will make appreciable contributions and the resulting equation is very difficult to handle.

III. Application to Experimental Work

(*a*) *Theoretical Predictions*

If the intrinsic breakdown in the low temperature region is assumed to be correctly described by the Frohlich-Paranjape (1956) theory and if at higher temperatures breakdown is thermal, then the breakdown strength of an alkali halide as a function of temperature would rise slowly with increasing temperature up to some transition temperature (which depends chiefly on the manner of application of the voltage, and the ionic conductivity at that temperature). It would thereafter drop rapidly in a way which is illustrated in Figure 3 for sodium chloride and potassium bromide for various voltage rise times and for ionic conductivities as given by Lehfeldt (1933).

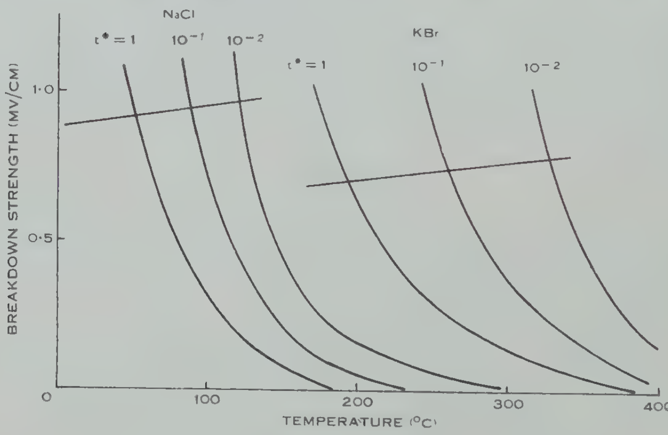


Fig. 3.—Theoretical breakdown strengths of sodium chloride and potassium bromide as a function of temperature for various breakdown times. The short almost horizontal lines are the Frohlich-Paranjape intrinsic theory, and the steeply falling curves " impulse thermal " from equation (**10**).

Application of the criterion (**11**) for critical rise time shows that $t^* \sim 10^{-1}$ sec marks the dividing line between impulse thermal and steady state thermal conditions for typical alkali halides ($C_v \sim 0 \cdot 25$ cal/cm^3 degC, $l \sim 10^{-1}$ cm, and $\varkappa \sim 0 \cdot 025$ cal/cm sec degC). Reference to Figure 3 shows that the values of t^* chosen are grouped about this critical value. This has been done since what are normally referred to as " d.c. tests " are probably as far removed from either approximation as it is possible to be, even though the voltage rise time is more of the order of 10 sec in " d.c. tests " than the 1/10 sec estimated above. The reason for this is that the evaporated metal films or smears of graphite used as electrodes by many workers are good electrical contacts but exceedingly poor thermal sinks. The value of l should then be greater than that estimated above and t^* accordingly greater also.

For the thicknesses used in experimental work the value of V^* from (**6a**) would give a breakdown field strength many orders of magnitude greater than that obtained from (**10**) for values of t^* shown in Figure 3. Thus, for thin crystals, the thermal breakdown strength does not converge slowly down to some limit as t^* becomes large, but reaches a minimum—further increase in t^* should then increase the breakdown field strength up to the limit given by (**6a**). It would seem, however, that careful control of experimental conditions would be needed to make the breakdown strength appreciably higher than the minimum.

The very great importance of the nature and concentration of impurities in the dielectric sample is now evident. The value of electrical conductivity which is of importance is that at the ambient temperature, and this will be in the so-called " structure sensitive region ". However, the conductivity in this region depends mainly on the concentration of divalent impurities (cf. Kelting and Witt 1949) so that from (**11**) we would expect the thermal breakdown strength to depend on the inverse square root of the divalent impurity concentration.

(*b*) *Comparison with Existing Experimental Work*

With regard to the apparent disagreement between Calderwood and Cooper (1953) on the one hand and Alger and von Hippel (1949) and Konorova and Sorokina (1957) on the other, a consistent explanation appears by examining what they regard as impulse breakdown. Thus in order to determine the impulse breakdown strength of a specimen Calderwood and Cooper applied a series of voltage pulses increasing by increments of 5 per cent. or less from about 70 per cent. of the estimated breakdown voltage. Depending on more exact details of the test this procedure may not be very different from some form of " d.c. test ". However, Alger and von Hippel (and presumably also Konorova and Sorokina) used a voltage increasing linearly up to breakdown. This is the type of voltage application assumed in the previous analysis, and is in fact the only reasonable way to apply voltage if thermal conditions are suspected to be of importance.

Turning then to the experimental results of Figure 1—they are of the same form and of the same order of magnitude as the theoretical results of Figure 3. Moreover, the discrepancies between the two sets of results in Figure 1 can be explained easily and naturally on the thermal breakdown theory. Since the type of voltage application is similar in the two cases, the experimental factors which may differ and are of importance are the impurity content of the crystals, and the thermal contact provided by the electrodes. Both of these factors could vary by orders of magnitude from one set of results to the other, and it would seem that Konorova and Sorokina used specimens of higher impurity content and electrodes which were poorer thermal sinks than those of Alger and von Hippel. In support of this contention we note in the first place that higher divalent impurity content in an alkali halide means reduced electronic mobility which, according to the Frohlich-Paranjape theory, means a higher breakdown strength in the intrinsic region. This higher divalent impurity content also means increased ionic conductivity and, since the thermal conductivity will not be greatly changed, will result in a rise of the ratio of the electrical to the thermal conductivity. Thermal breakdown, if it occurs, should then occur at a lower temperature than otherwise. Both of these predicted consequences of higher impurity content are observed in the results of Konorova and Sorokina as shown in Figure 1. Secondly, the peculiar " hump " in the results of these authors can be explained if their electrodes were relatively poor thermal sinks. (Similar " humps " in the breakdown characteristics of potassium bromide were also observed by Buehl and von Hippel (1939) and Austen and Whitehead (1940).) In these circumstances Joule heat generated inside the specimen would cause a rise in temperature that may not be sufficient to cause thermal breakdown for the time for which the voltage is applied, but yet sufficient to require a higher breakdown strength of intrinsic nature. Since d.c. tests are commonly complete in time of order of 100 sec and about 100 degC rise in temperature would be required to appreciably alter the intrinsic breakdown strength, the field would need to produce a temperature rise of order of 1 degC per sec. With field strengths of order of 10^6 V/cm and specific heat of order of 1 joule/cm³ the required electrical resistivity is about $10^{12}\,\Omega$ cm, which is of the order of the resistivity of potassium bromide for the temperature at which the " hump " begins to appear.

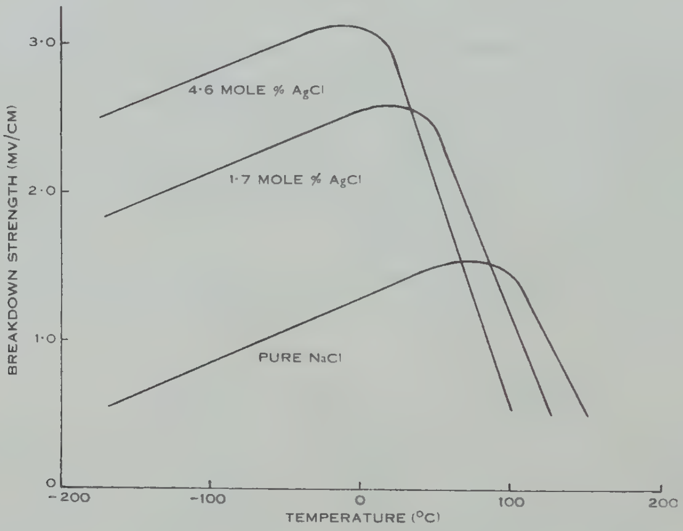


Fig. 4.—The dielectric breakdown strength of sodium chloride with various additions of silver chloride as a function of temperature (von Hippel and Lee).

Further support for this explanation for the experimental results on potassium bromide comes from the work of von Hippel and Lee (1941) on crystals of sodium chloride to which silver chloride had been added. Their results are shown in Figure 4 and tend to bear out the interpretation given above. It is also interesting to note that in the region of presumed thermal breakdown the breakdown strength is lower for the crystals of higher electrical conductivity.

(c) *Suggestions for Future Work*

In view of what has been said above three main suggestions concerning future work emerge. They are:

(i) That the impurity concentration (especially of divalent ions) may be an all important factor. It is even conceivable that unknown and random concentrations of divalent impurities may account for much of the higherto unexplained scatter of experimental results. It would thus seem to be important that experimental work be done on crystals of known impurity concentration.

(ii) Since thermal conditions of test would play an important part in thermal breakdown then these conditions should be stated and controlled—preferably by using electrodes which are good thermal sinks.

(iii) Since breakdown is clearly intrinsic at very low temperatures and clearly thermal at very high temperatures, experimental study in the transition region may always present grave difficulties. In view of the extreme sensitivity of breakdown strength to some parameters, intrinsic and thermal breakdown may occur in different samples of the one substance at the same temperature and with the voltage applied in the same way. A thorough study of both ends (i.e. very high and very low temperatures) may therefore be profitable before endeavouring to sort out details in the transition region.

IV. References

Alger, R. S., and von Hippel, A. (1949).—*Phys. Rev.* **76**: 127.
Austen, E. W., and Whitehead, S. (1940).—*Proc. Roy. Soc.* A **176**: 33.
Buehl, R. C., and von Hippel, A. (1939).—*Phys. Rev.* **56**: 941.
Calderwood, J. H., and Cooper, R. (1953).—*Proc. Phys. Soc. Lond.* B **66**: 73.
Franz, W. (1952).—*Z. Phys.* **132**: 285.
Franz, W. (1956).—" Encyclopedia of Physics." Vol. 17. p. 155. (Springer: Berlin.)
Frohlich, H., and Paranjape, B. V. (1956).—*Proc. Phys. Soc. Lond.* B **69**: 21.
von Hippel, A., and Lee, G. M. (1941).—*Phys. Rev.* **59**: 824.
Kelting, H., and Witt, H. (1949).—*Z. Phys.* **126**: 697.
Klemens, P. G. (1958).—*Solid State Phys.* **7**: 1.
Konorova, E. A., and Sorokina, L. A. (1957).—*J. Expt. Theor. Phys.* **32**: 143.
Lehfeldt, W. (1933).—*Z. Phys.* **85**: 717.
Mott, N. F., and Gurney, R. W. (1948).—" Electronic Processes in Ionic Crystals." 2nd Ed. (Oxford Univ. Press.)
Vermeer, J. (1956).—*Physica* **22**: 1247, 1257, 1269.
Whitehead, S. (1950).—" Dielectric Breakdown of Solids." (Oxford Univ. Press.)

THE INTERACTION OF DISLOCATIONS WITH BOUNDARIES AND SURFACE FILMS

By A. K. Head*

[*Manuscript received September 25, 1959*]

Summary

A digital computer programme for finding the equilibrium positions of a group of dislocations is applied to two problems of a different type from those considered previously. These are:

(i) An infinite solid has different shear moduli on either side of a plane boundary. On the side with lower modulus a group of dislocations are on a slip plane which is perpendicular to the boundary. They are forced towards the boundary by an applied stress but repelled from the boundary by the change in modulus.

(ii) On the plane face of a semi-infinite solid is a thin film with shear modulus greater than that of the bulk. A group of dislocations are on a slip plane which is perpendicular to the film. The applied stress necessary to force the leading dislocation into the film is found in two particular cases which correspond to a rhodium film on silver or to the oxide film on aluminium.

I. Introduction

From their experiments on the deformation of thin films of face-centred cubic metals on silver or palladium substrates, Brame and Evans (1958) concluded that there could be two barriers preventing dislocations passing into the film from the substrate. The first is the dislocation array at the interface caused by the lattice misfit of film and substrate and the second is the difference in elastic constants of film and substrate. The latter effect is quite a complicated dislocation interaction problem but is within the range of problems for which numerical solutions can be obtained using an electronic computer (Head 1959).

It has been shown previously (Head 1953) that if there is a surface film whose shear modulus is greater than that of the substrate then there is a long-range attraction and a short-range repulsion on a dislocation in the substrate. This force can be described by an infinite set of image dislocations. The equilibrium position of a single dislocation was calculated numerically as a function of the ratio of shear moduli.

In Section III the following problem is analysed. Consider n dislocations on the same slip plane in a substrate of shear modulus μ_1 with a surface film of shear modulus μ_2. A shear stress σ is applied which moves the dislocation array towards the film-substrate interface, and the value of σ necessary to overcome the repulsion due to the film is calculated. As a preliminary, the case of a very thick film is considered in Section II. As usual, the image representation is exact for screw dislocations and a good approximation for edge dislocations.

---

* Division of Tribophysics, C.S.I.R.O., University of Melbourne.

II. Interaction with Thick Film

Consider an infinite elastic medium of shear modulus μ_1 for $x>0$ and modulus μ_2 for $x<0$. Then the stress field of a unit dislocation at $(x,0)$ can be written as the sum of two parts. The first is the stress field of the unit dislocation in a homogeneous medium and the second is the stress field of an image dislocation of strength $K=(\mu_2-\mu_1)/(\mu_2+\mu_1)$ at the image point $(-x,0)$. If $\mu_2>\mu_1$, the dislocation and image are of the same sign, giving a repulsion of the dislocation from the interface.

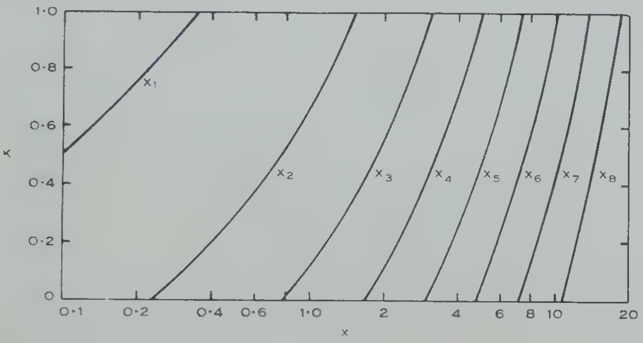


Fig. 1.—Positions of dislocations as a function of $K=(\mu_2-\mu_1)/(\mu_2+\mu_1)$ for $n=8$. Unit of length $=\mu_1 b/2\pi\sigma$.

Suppose there are n dislocations on the x axis at $x_1, x_2, \ldots, x_n$ and there is an applied stress σ forcing these dislocations towards the interface. Then the equilibrium positions of the dislocations are given by

$$\sum_{\substack{i=1\\ \neq j}}^{n} \frac{A}{x_j-x_i}+K\sum_{i=1}^{n} \frac{A}{x_j+x_i}-\sigma=0, \quad j=1, \ldots, n, \qquad \ldots\ldots (1)$$

where $A=\mu_1 b/2\pi$.

As a preliminary to finding solutions of these equations it is convenient to take the unit of length as A/σ. The equations (1) can then be written in the non-dimensional form

$$\sum_{\substack{i=1\\ \neq j}}^{n} \frac{1}{x_j-x_i}+K\sum_{i=1}^{n} \frac{1}{x_j+x_i}-1=0, \quad j=1, \ldots, n. \qquad \ldots\ldots (2)$$

As a typical example Figure 1 shows, for $n=8$, the non-dimensional positions of the eight dislocations as K varies from zero to unity. A reasonable summary of the properties of this type of dislocation array is given by Figures 2 and 3. In Figure 2 is shown the position x_n of the dislocation furthest from the interface as a function of n and K, and in Figure 3 the position x_1 of the dislocation closest to the interface.

It is interesting to compare this dislocation array with the well-known case of a pile-up against a locked dislocation (Eshelby, Frank, and Nabarro 1951). In both the cases the positions of the dislocations vary inversely with the applied stress but the dependence on n, the number of dislocations, is different. Eshelby, Frank, and Nabarro showed that the distance of the leading dislocation from the barrier (the locked dislocation) was approximately proportional to n^{-1}, whereas from Figure 3 it will be seen that the corresponding distance from the barrier (the interface) is less dependent on the number of dislocations, being approximately proportional to $n^{-\frac{1}{2}}$.

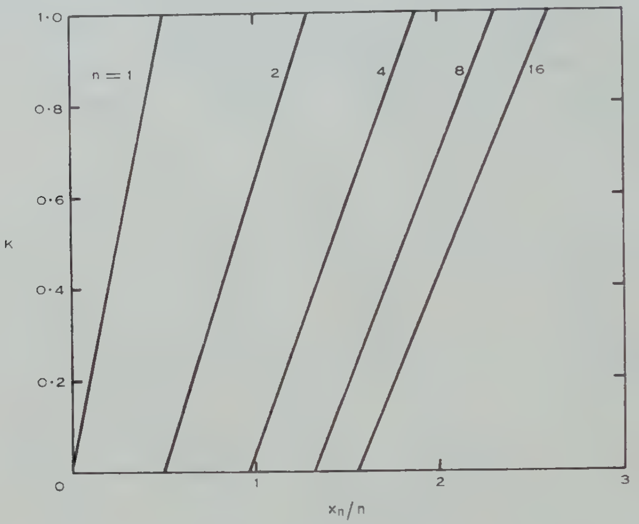


Fig. 2.—Position of outer dislocation as a function of K and n. Unit of length=$\mu_1 b/2\pi\sigma$.

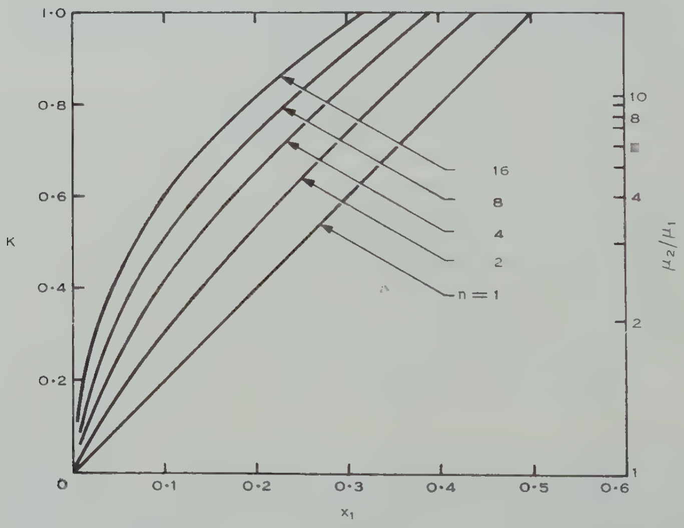


Fig. 3.—Position of inner dislocation as a function of K and n. Unit of length=$\mu_1 b/2\pi\sigma$.

III. INTERACTION WITH THIN FILM

Suppose now that the situation of Section II is changed so that the medium of shear modulus μ_2, which is to the left of the plane $x=0$, is of thickness d. Then it has been shown (Head 1953) that a dislocation in the region $x>0$ can be considered as generating an infinite array of image dislocations and that the conditions of equilibrium of n dislocations arranged along the x-axis are

$$\sum_{\substack{i=1\\ \neq j}}^{n} \frac{A}{x_j-x_i}+\sum_{i=1}^{n}\left\{\frac{KA}{x_j+x_i}-\sum_{m=1}^{\infty}\frac{(1-K^2)K^{m-1}A}{x_j+x_i+2md}\right\}-\sigma=0, \quad j=1, \ldots, n, \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

where $A=\mu_1 b/2\pi$, $K=(\mu_2-\mu_1)/(\mu_2+\mu_1)$, and σ is the applied stress.

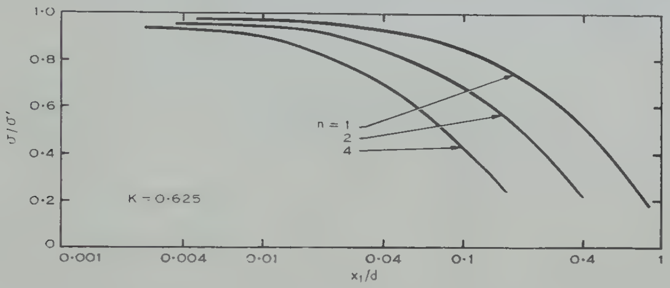


Fig. 4.—Position of inner dislocation as a function of σ and n for $K=0\cdot625$.

A summary of some solutions of these equations is given in Figure 4 for $K=0\cdot625$ and in Figure 5 for $K=0\cdot5$, these two cases representing a rhodium film on a silver substrate and an aluminium oxide film on aluminium. These figures give the position of the leading dislocation in the group as a function of the applied stress in the following manner. The distance of the dislocation from the film-substrate interface is given as a fraction of the film thickness d, and the stress σ necessary to bring the dislocation to this position is given as a fraction of the stress σ' which would be needed if the film were infinitely thick (Section II). Due to the complexity of the equations (3) it was not practicable to compute beyond $n=4$. However, Figures 4 and 5 suggest that the curves

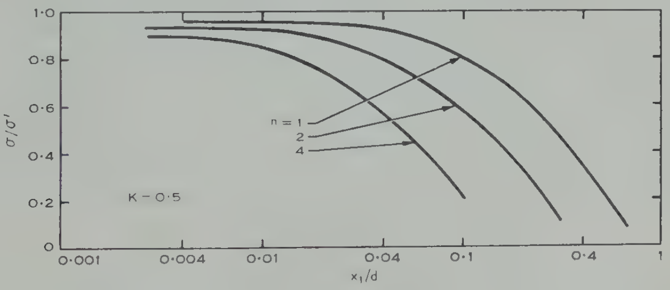


Fig. 5.—Position of inner dislocation as a function of σ and n for $K=0\cdot5$.

are a similar shape and the effect of doubling the number of dislocations is to shift the curves a constant distance to the left. This approximation will be used to extrapolate to larger n.

We now apply these calculations to the experimental observation of Brame and Evans that a rhodium film mounted on a silver substrate is brittle, with the implication that it is not possible to inject dislocations into the film from the substrate. Although it would require an infinite stress to move an ideal dislocation right up to the film-substrate interface against the image repulsion, the repulsion on a real dislocation will saturate when the core of the dislocation reaches the boundary. Suppose then we consider the stress necessary to bring the leading dislocation to a distance of $2b$ from the interface. Taking the thickness of Brame and Evans's rhodium film as $200b$ then from Figures 3 and 4 we can derive Table 1.

TABLE 1

STRESS σ REQUIRED TO FORCE n DISLOCATIONS INTO RHODIUM FILM FROM SILVER SUBSTRATE

n	1	2	4	8	16	32	64
$\frac{\sigma}{\mu_1}$	$\frac{1}{40}$	$\frac{1}{55}$	$\frac{1}{80}$	$\frac{1}{110}$	$\frac{1}{180}$	$\frac{1}{450}$	0

For $n \geqslant 8$ this table involves extrapolation of the curves of Figure 4. However, a check on the last entry can be obtained as follows. When there is no applied stress then the situation is very similar to that considered in Section 5 of Head (1959). This was the case of n dislocations attracted to a free surface but being held up by a locked dislocation. If the distance of this locked dislocation from the free surface is d then it was shown that the stress field at the locked dislocation, due to the mobile dislocations and their images, was $\frac{1}{2}n\mu b/2\pi d$. Now in the case of the surface film with no applied stress only the leading dislocation is close to the film, and the other dislocations are attracted as if by a free surface. Thus the stress forcing the leading dislocation towards the film will be approximately $\frac{1}{2}n\mu_1 b/2\pi d$ and if it is close to the film then the approximate repulsion from the film will be that due to an image dislocation of strength K. The equilibrium distance x_1 of the leading dislocation from the interface will thus be given approximately by

$$\left.\begin{aligned} \frac{n}{2}\,\frac{\mu_1 b}{2\pi d} &= \frac{K\mu_1 b}{2\pi(2x_1)}, \\ \text{that is,} \qquad n x_1 &= dK. \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

With $x_1 = 2b$, $d = 200b$, $K = 0\cdot625$, this gives $n = 62\cdot5$, which is to be compared with the value $n = 64$ in Table 1. This suggests that the extrapolation involved in Table 1 is quite reasonable.

From Table 1 it will be seen that the applied stress necessary to overcome the elastic repulsion from the film is large for a small number of dislocations and that it only falls slowly as the number of dislocations is increased up to 32 but that it then falls suddenly to zero for $n \sim 64$. This suggests that the number of dislocations on a slip plane in the silver substrate is probably less than 32.

IV. REFERENCES

BRAME, D. R., and EVANS, T. (1958).—*Phil. Mag.* **3**: 971.

ESHELBY, J. D., FRANK, F. C., and NABARRO, F. R. N. (1951).—*Phil. Mag.* **42**: 351.

HEAD, A. K. (1953).—*Phil. Mag.* **44**: 92.

HEAD, A. K. (1959).—*Phil. Mag.* **4**: 295.

THE INFLUENCE OF TEMPERATURE AND STRAIN RATE ON THE FLOW STRESS OF MAGNESIUM SINGLE CRYSTALS

By Z. S. Basinski*

[*Manuscript received September 25, 1959*]

Summary

The temperature and strain rate-dependence of the flow stress of magnesium single crystals was measured in the temperature range 1·45–248 °K. Some incidental observations were also made on the form of the stress-strain curves and the appearance of slip lines. The deformation characteristics show close similarity with those of the face-centred cubic metals.

I. Introduction

Although a considerable amount of work, both experimental and theoretical, has been done on the mechanism of strain hardening in single crystals of face-centred cubic metals, comparatively little is known about the deformation behaviour of close-packed hexagonal metals. Especially lacking is information on the deformation characteristics at low temperatures, where the complicating effects of annealing during glide are absent.

The only available stress-strain curves at low temperatures are of magnesium (Schmid and Siebel 1931) and of cadmium (Boas and Schmid 1930) single crystals, which show a region of low hardening (rate of hardening roughly corresponding to stage I in face-centred cubic crystals) up to strains of the order of 1·5. This is then followed by much more rapid hardening. Unfortunately zinc, which has been investigated more extensively at low temperatures, fractures quite early in the deformation.

No detailed theory of strain hardening of hexagonal single crystals has been proposed, but the mechanism is thought to be analogous to that in stage I of the deformation of face-centred cubic metals (Mott 1957; Seeger 1957), the hardening being ascribed to elastic interactions between glide dislocations and piled-up groups of other dislocations lying on parallel planes. Recent work on face-centred cubic metals (Basinski 1959; Hirsch, personal communication 1959) casts serious doubt on the validity of an elastic interaction mechanism, and hardening, at any rate in stages II and III, has been shown to be more probably a result of accumulation of a forest of dislocations cutting the glide plane.

A distinction between elastic interaction and any mechanism involving short-range obstacles (e.g. forest hardening) can be made by measurements of the influence of temperature and strain rate on the reversible flow stress (Cottrell and Stokes 1955). The activation energy associated with overcoming long-range obstacles would be outside the range of thermal fluctuations, the flow stress would be strain-rate independent, and dependent on temperature only through

---

* Division of Pure Physics, National Research Council, Ottawa, Canada.

the elastic moduli. In contrast, where hardening is due to short-range obstacles, the flow stress will depend on the strain rate, and its temperature dependence will obviously be more than that due to the elastic moduli. For thermally activated glide through randomly distributed short-range obstacles whose interaction with glide dislocations is proportional to the elastic modulus, the influence of temperature and strain rate on the flow stress is related by (Basinski 1959)

$$\ln\,(NkTAb\nu/\dot{\varepsilon}) = -\left[1+\frac{\left(\partial\dfrac{\sigma/\mu}{\sigma_0/\mu_0}/\partial T\right)_{\dot{\varepsilon}}}{\dfrac{\sigma/\mu}{\sigma_0/\mu_0}\,\dfrac{1}{T}\left(\dfrac{\partial \ln \sigma}{\partial \ln \dot{\varepsilon}}\right)_T}\right]\Bigg/\left(1-\frac{T\partial \ln \mu}{\partial T}\right), \quad \ldots \quad (\mathbf{1})$$

where $N\mathrm{d}U$ is the number of activation sites having activation energy between U and $(U+\mathrm{d}U)$, A is the area swept by a dislocation after passing an obstacle, ν is the vibration frequency of a dislocation against an obstacle, $\dot{\varepsilon}$ is the strain rate, σ is the flow stress, and μ is the elastic modulus.

In face-centred cubic metal crystals, the change of stress resulting from change in temperature or strain rate, increases remarkably linearly with total stress. This linear relation, the Cottrell-Stokes law, can be well explained in terms of a forest hardening mechanism.

The object of the present work was to determine the extent to which short-range obstacles are responsible for hardening in hexagonal metals.

II. Experimental Details

Specimens were prepared from high purity magnesium (99·98 per cent.) supplied by Dominion Magnesium Company Limited. Single crystals of length 20–25 cm and diameter 2·5 mm were grown from the melt in split moulds of spectroscopically pure graphite in a helium atmosphere. These were then cut into lengths of about 3·5 cm in an acid cutting machine, and the specimens numbered consecutively A, B, C, . . . from the end first to solidify. The orientation of the first and last specimen from each crystal was determined by the standard rotation method (Schmid and Boas 1950). If these two orientations were different, all specimens from that crystal were rejected. The orientations of specimens used are shown in Figure 1.

All tests were carried out in a hard beam resistance strain gauge tensile testing machine (Basinski 1959; see also White 1959). The specimens were mounted in split-jaw grips with an exposed length of 2·5 cm. Immediately before mounting, each crystal was chemically polished in concentrated nitric acid, and its diameter measured. This polishing treatment produced a surface sufficiently good for subsequent metallographic analysis. With the exception of one run at liquid nitrogen temperature, the specimens were sealed in the cryostat and evacuated overnight. Before testing, about half an atmosphere of helium exchange gas was admitted to ensure good thermal contact between the specimen and the cryostat. The temperatures used were 1·45, 4·2, 10·25, 13·4, 20·35, 28·8, 37·8, 46, 63, 77·8, 90, 100, 127, 163, 200, and 248 °K, and were measured by a calibrated platinum resistance thermometer, with the exception of 1·45 °K, for which a helium vapour-pressure thermometer was used. Some runs were carried out at constant temperature at two strain rates, corresponding to crosshead speeds of $1\cdot25\times10^{-4}$ and $1\cdot25\times10^{-3}$ cm/sec and in several cases between $1\cdot25\times10^{-5}$ and $1\cdot25\times10^{-4}$ cm/sec. In other runs, the temperature as well as the strain rate was changed repeatedly during the deformation. For temperatures above 78 °K the changes were made between a temperature T and 78 °K, and for all other temperatures, between T and 4·2 °K.

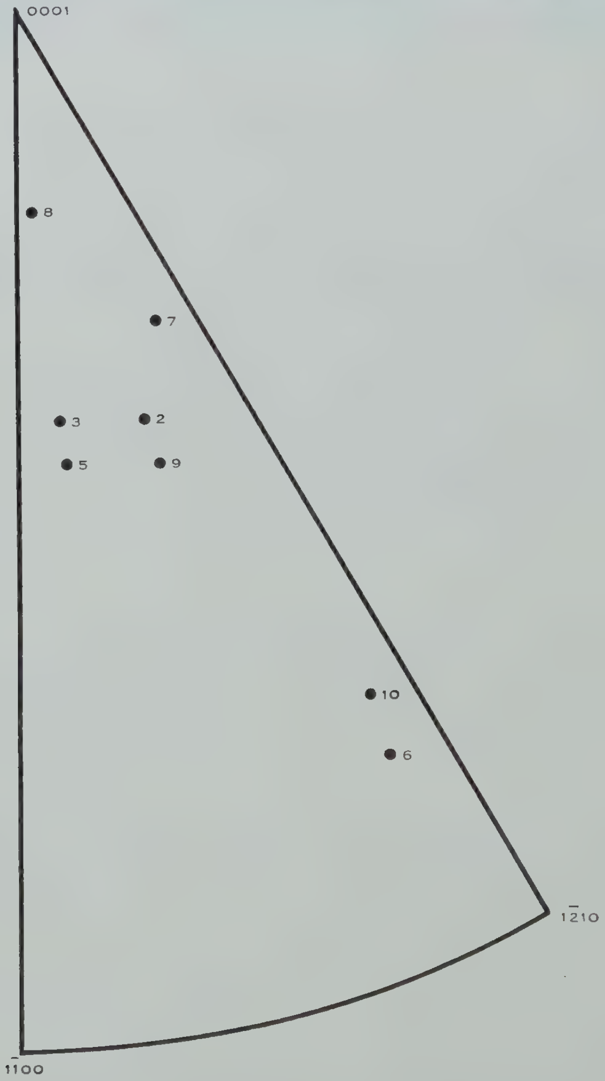


Fig. 1.—Orientations of crystals tested.

No unloading yield-point effects (Haasen and Kelly 1957 ; Makin 1958), usually troublesome in this type of investigation for face-centred cubic metals, were observed for magnesium. Nevertheless, the unloading during the temperature changes was limited to about 10 per cent. of the flow stress at temperatures above that of liquid nitrogen, and 2–3 per cent. below this temperature. The load-elongation curve was recorded at a sensitivity of 50 g/in. for ten zero suppressions, i.e. until a load of 4 kg was reached. The sensitivity was then halved for loads up to 8 kg, and subsequently halved again. The strain was obtained from the crosshead motion. The final orientation of the specimen was checked in several cases with that predicted from the strain, the agreement always proved satisfactory.

The resolved shear stress and strain were calculated from the load-elongation curves by using the standard formulae of Schmid and Boas (1950). After fracture, the surface of the specimen was usually examined under an optical microscope. Slip markings formed on one of the specimens were examined at various stages of the deformation by removing the previous markings by repolishing the surface (Blewitt, Coltman, and Redman 1955).

III. EXPERIMENTAL RESULTS

(a) *Temperature and Strain-rate Dependence of the Flow Stress*

The general appearance of the stress-strain curves following change in either strain rate or temperature was similar to that previously observed in aluminium, copper, and silver (Basinski 1959). The only notable exception was the absence of the pseudo yield points which were observed following decrease in temperature for certain temperature ranges in the face-centred cubic metals. This most probably results from the absence of the unloading yield-point phenomenon in magnesium. On the other hand, similarly to the behaviour in face-centred cubic metals, the slope of the curve was usually somewhat smaller than average immediately after a steep increase in stress (caused either by increased strain rate or decreased temperature) and somewhat larger than average after the flow stress had been sharply decreased.

The method of interpretation of the stress changes adopted in the present work was the same as that used for the face-centred cubic metals, i.e. the values of stress corresponding to the two temperatures or strain rates immediately adjacent on the strain axis were used. This, it was hoped, would give the flow stress corresponding to the same dislocation configuration.

In some specimens regular oscillations in the flow stress were observed in the earlier stages of deformation, especially at the lower strain rate. No correlation of the appearance of this waviness was found with temperature—it was observed at 4·2 °K and at room temperature. Correlation with orientation was not rigid, some specimens with a given orientation exhibited this behaviour, while others did not. However, for certain orientations, i.e. series 6 and 10, whose tension axes lay near the basal plane, this behaviour was completely absent. The oscillations never persisted beyond 70 per cent. deformation. Similar oscillations were shown in creep tests on magnesium single crystals (Conrad and Robertson 1957). Since it was difficult to decide how to assign the change in stress resulting from a change in strain rate in the cases where these oscillations occurred, the strain-rate sensitivity was not evaluated here. In all cases the increase in stress was taken in the upward direction, i.e. following increase in strain rate or decrease in temperature.

At temperatures below about 46 °K, the Cottrell-Stokes law was obeyed. Figure 2 shows plots for the change in resolved shear stress $(\Delta\sigma)_T$ produced by increase in strain rate by a factor of 10, plotted against the total shear stress.

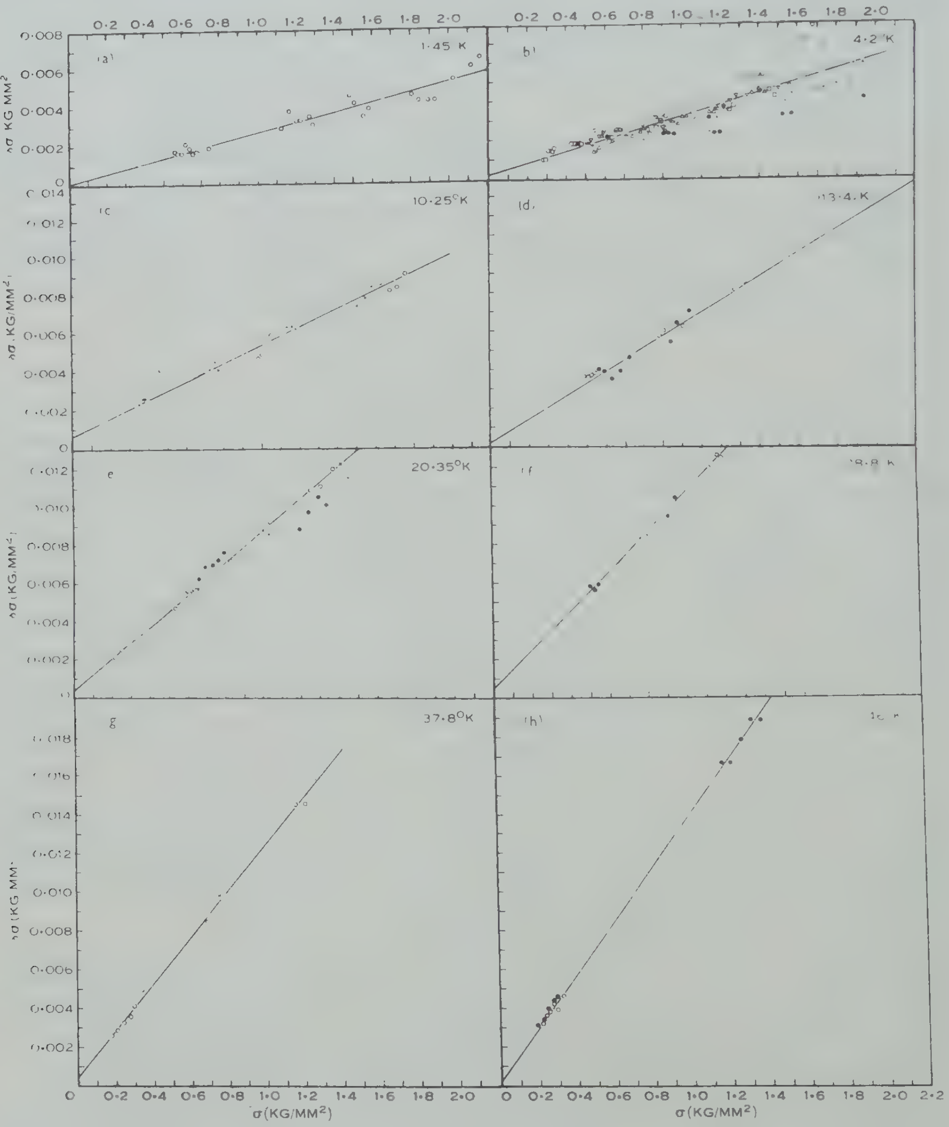


Fig. 2.—Change in resolved shear stress produced by increasing the strain rate by a factor of 10 plotted against the shear stress corresponding to the higher strain rate, at low temperatures.

$(\Delta\sigma)_T$ increases linearly with σ, this rate of increase being dependent on temperature. Although the curves do not extrapolate to zero at $\sigma=0$, the intercept here is very low, being of the order of the experimental scatter. The results of temperature changes follow essentially the same pattern.

The relation between $(\Delta\sigma)_T$ and σ becomes somewhat more complex at higher temperatures, as shown in Figure 3. Starting with data at 63 °K up to 90 °K, $(\Delta\sigma)_T$ increases linearly at first and then shows a deviation at higher stresses, its value being somewhat lower than those obtained by extrapolation.

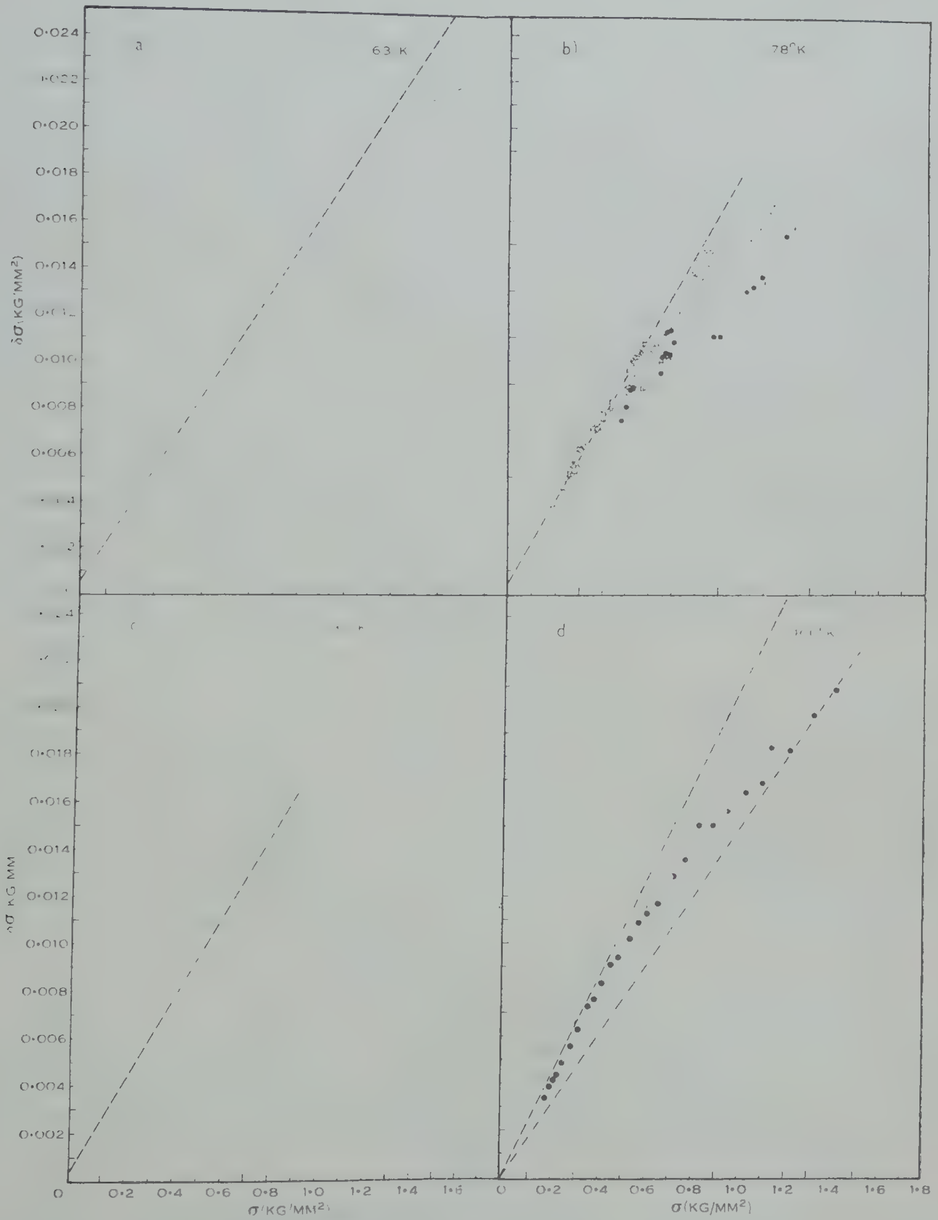


Fig. 3 (*a*)–(*d*).—Same as Figure 2, at temperatures above 46 °K. Different symbols signify different specimens, except in (*b*), where they signify different predeformation temperatures: ▽, 4·2 °K; ○, 78 °K; ×, 90 °K; △, 127 °K; +, 163 °K; ●, 200 °K; □, 248 °K.

It is interesting to note that this deviation appears to be a function of the temperature of prestrain, increasing with prestrain temperature. In Figure 3 (*b*) the points were identified by the temperature which immediately preceded deformation at 78 °K. In the region of linear increase of $(\Delta\sigma)_T$ with σ the results do not appear to be a function of prestrain temperature.

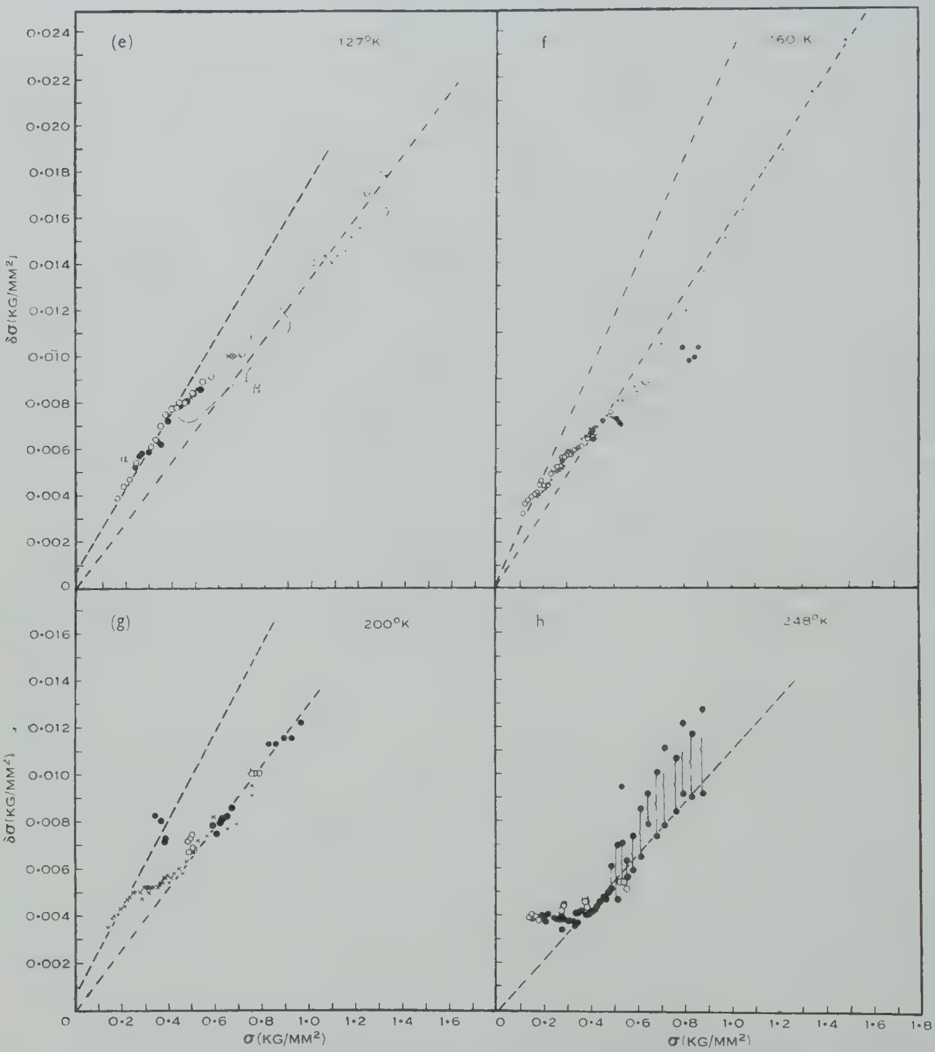


Fig. 3 (*e*)–(*h*).—In (*h*) at higher stresses at 248 °K the stress-strain curves following change of strain rate were rounded. The points joined by vertical lines represent the possible limits of $(\Delta\sigma)_T$.

At still higher temperatures, curves of the type shown in Figures 3 (*d*)–3 (*h*) were obtained. In Figure 3 (*e*), the decrease in the slope of the $(\Delta\sigma)_T/\sigma$ curve (portion β) is followed at higher stresses by an increase in slope (portion γ of the same figure). The two portions α and γ seem to extrapolate roughly through the origin of coordinates. The transition between α and γ moves toward lower stresses with increasing temperature until about 248 °K, when portion α is not observed. A curve similar to that in magnesium at 248 °K was observed for a cadmium single crystal at room temperature by Christian and his co-workers (Christian, personal communication 1959).

These results could suggest that hardening in magnesium is governed by two similar kinds of obstacle, one characteristic of low stresses and low temperatures, the other characteristic of higher stresses and temperatures, the Cottrell-Stokes law being obeyed by each. The complexity of the $(\Delta\sigma)_T/\sigma$ curves arises from the combined operation of the two mechanisms. It is quite probable that this type of behaviour also occurs in the face-centred cubic metals. The data available there (Adams and Cottrell 1955 ; Cottrell and Stokes 1955 ; Seeger *et al.* 1957 ; Basinski 1959) mainly correspond to the higher stresses reached during stages II and III of the deformation. They would then correspond mainly to the part γ of the magnesium curve except at very low temperatures. The sign of the deviation from the Cottrell-Stokes law in face-centred cubic metals at low stresses appears consistent with the observations on magnesium.

Stress changes produced by variation of temperature followed a pattern similar to that observed for the strain-rate changes. At temperatures below that of liquid nitrogen the Cottrell-Stokes law was obeyed quite well. At higher temperatures, the increase in the resolved shear stress with the total shear stress was somewhat slower than that indicated by the Cottrell-Stokes law. Unfortunately, the number of points available was not sufficient to determine unambiguously the $\Delta\sigma/\sigma$ relation ; however, this behaviour is consistent with the observations on the strain-rate changes.

Figure 4 shows the ratio of the flow stress of magnesium at a temperature T to that at $4\cdot2$ °K determined by two methods. The data above 78 °K were normalized by assuming the validity of the relation $\sigma_T/\sigma_{4\cdot2}=\sigma_T/\sigma_{78}\cdot\sigma_{78}/\sigma_{4\cdot2}$ (Cottrell and Stokes 1955). The lower curve with points marked by + was obtained by arbitrarily setting $\sigma=1\cdot2$ kg/mm^2 and taking the flow stress ratio at this stress. The slightly higher curve, marked with □, was obtained by taking the slope of the best straight line drawn through the plots of $\Delta\sigma/\sigma$. The two upper curves give the values after correction for the temperature dependence of the elastic modulus (C_{44} taken from the work of Slutsky and Garland 1957). The numerical values of $\sigma_T/\sigma_{4\cdot2}$ are listed in Table 1.

The values of the strain-rate sensitivity at different temperatures are given in Table 2 and Figure 5. Because of the difficulties of interpretation in this case, the values of the strain-rate sensitivity listed here were obtained from (*a*) the slope of the portion α of the $\Delta\sigma/\sigma$ curve, and (*b*) from the values at $\sigma=1\cdot2$ kg/mm^2, which at high temperatures corresponds approximately to portion γ.

(*b*) *The Relation between the Temperature and Strain-rate Dependence of the Flow Stress*

If the reversible change of flow stress with both temperature and strain rate is a result of thermal activation of dislocations over obstacles, it follows from equation (1) that the strain-rate sensitivity, $(1/T)\mathrm{d}\ln\sigma/\mathrm{d}\ln\dot{\varepsilon}$, should be approximately proportional to the fractional rate of change of flow stress with temperature. Since none of the terms under the logarithm sign, with the exception of T, should sharply depend on temperature, the value of the logarithm should be approximately constant since the variation in T should not affect the value of $\ln NkTAb\nu/\dot{\varepsilon}$ except at very low temperatures.

The value of $\ln NkTAb\nu/\dot{\varepsilon}$ obtained from the temperature and strain-rate data by using equation (**1**) at temperatures above 10 °K is approximately 17–18. The corresponding value for face-centred cubic metals falls in the range 13·5–15 (Basinski 1959). In view of the scatter of the experimental results and the somewhat arbitrary choice of the elastic moduli, this discrepancy is not considered serious, and the similarity of the values suggests that the deformation in magnesium is controlled by a thermally activated process similar to that in the face-centred cubic metals.

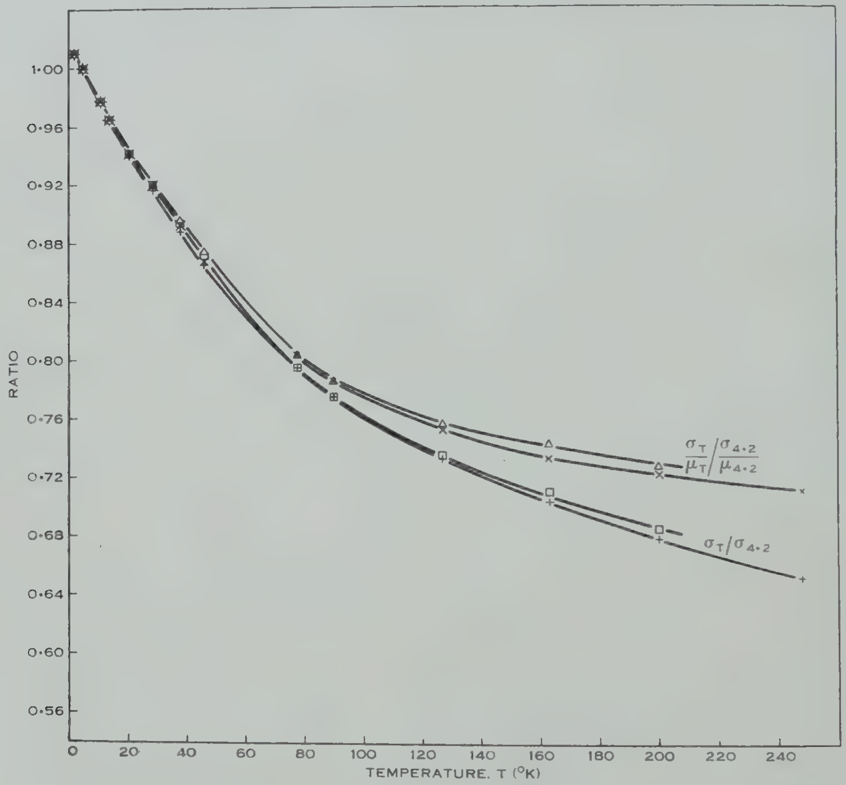


Fig. 4.—The ratio of flow stresses at temperature T to that at 4·2 °K. +, Value of the ratio taken at $\sigma = 1{\cdot}2\ \mathrm{kg/mm^2}$; □, obtained from the slopes of the $\Delta\sigma/\sigma$ curves.

At 4·2 °K and especially at 1·45 °K, the strain-rate sensitivity has values higher than those expected from the temperature variation of the flow stress. A similar observation was made for aluminium. The deviation is too large to be explained by the decrease with T of the value of $\ln NkTAb\nu/\dot{\varepsilon}$, and therefore

TABLE 1

FLOW STRESS AT TEMPERATURE T, RELATIVE TO THAT AT 4·2 °K, TAKEN AT $\sigma=1\cdot2$ KG/MM²

Temperature (°K)	$\sigma_T/\sigma_{4\cdot2}$	$\dfrac{\sigma_T}{\mu_T}\Big/\dfrac{\sigma_{4\cdot2}}{\mu_{4\cdot2}}$
1·45	1·009	1·009
10·25	0·97584	0·97593
13·4	0·96350	0·96278
20·25	0·93884	0·93949
28·8	0·91467	0·91604
37·8	0·88834	0·89083
46·0	0·86550	0·86897
77·8	0·79375	0·80290
90	0·77404	0·78542
127	0·73223	0·75239
163	0·70313	0·73366
200	0·67965	0·72264
248	0·65286	0·71187

TABLE 2

THE VALUES OF THE STRAIN-RATE SENSITIVITY AT DIFFERENT TEMPERATURES

Temperature (°K)	$\dfrac{1}{T}\left(\dfrac{\partial \ln \sigma}{\partial \ln \dot{\varepsilon}}\right)_T \times 10^5$	
1·45	77·7*	80·8†
4·2	30·4	33·0
10·25	20·2	22·4
13·4	20·1	20·4
20·35	17·9	18·6
28·8	15·2	15·8
37·8	13·9	14·2
46·0	13·0	13·3
63·0	10·7	10·7
77·8	9·8	7·2
90·0	8·2	—
100	8·8	6·4
127	5·7	4·6
160	4·1	4·2
200	2·8	2·8
248	2·1	2·1

* Values in this column obtained from the initial slope of $(\Delta\sigma)_T/\sigma$ curves.

† Values in this column correspond to $\sigma=1\cdot2$ kg/mm².

most probably is a manifestation of zero-point energy activated flow (Glen 1956 ; Mott 1956). The zero-point energy fluctuations, if this explanation is correct, would be of about the same magnitude as thermal fluctuations at 2–3 °K, in agreement with Mott's calculation.

(c) *Appearance of the Stress-Strain Curves*

Typical stress-strain curves obtained in the present work are shown in Figures 6 and 7. In practically all cases, the curves were S-shaped, showing a minimum rate of strain hardening at intermediate strains. An exception to this was provided by the behaviour of specimens of orientation 6, where at liquid helium temperature the curves were almost linear with a rate of strain hardening of about 750 g/mm².

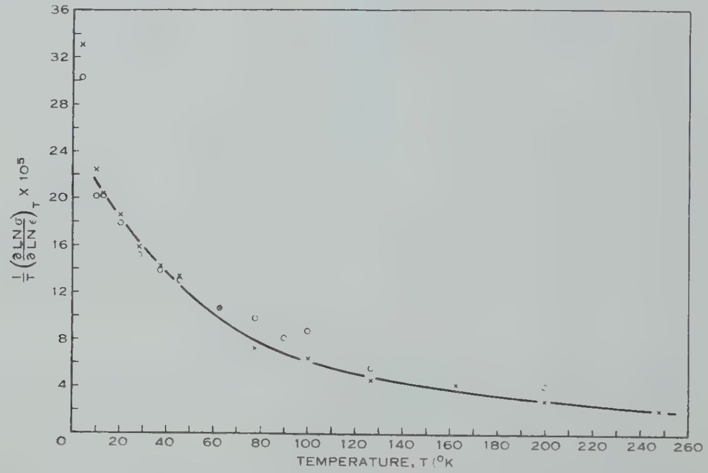


Fig. 5.—Temperature dependence of the strain-rate sensitivity. ×, Taken at $\sigma = 1 \cdot 2$ kg/mm²; ○, obtained from the initial slopes of the $\Delta\sigma/\sigma$ curves.

It will be convenient to divide the strain-hardening behaviour into three stages, A, B, and C ; the relative proportions of the three stages vary with temperature. Stage A at the beginning of deformation is characterized by a rate of strain hardening of about 250–400 g/mm², its value being approximately independent of temperature at temperatures below about 200 °K. At higher strains the rate of strain hardening decreases (stage B), and is followed by another region of increase in the strain-hardening rate (stage C). The amount of strain occupied by stage B varied sharply with temperature from zero at temperatures in the helium range to occupying most of the stress-strain curve at temperatures above about 250 °K. Stage C was characterized by an increasing rate of strain hardening which usually, at temperatures below 200 °K, reached a value of 2–3 kg/mm² before fracture. The stress-strain curves published by Schmid and Siebel (1931) for magnesium and by Boas and Schmid (1930) for cadmium show this behaviour quite strikingly.

In the experiments in which the temperature was changed during the deformation, the rate of strain hardening was appreciably affected by temperature

FLOW STRESS OF MAGNESIUM SINGLE CRYSTALS

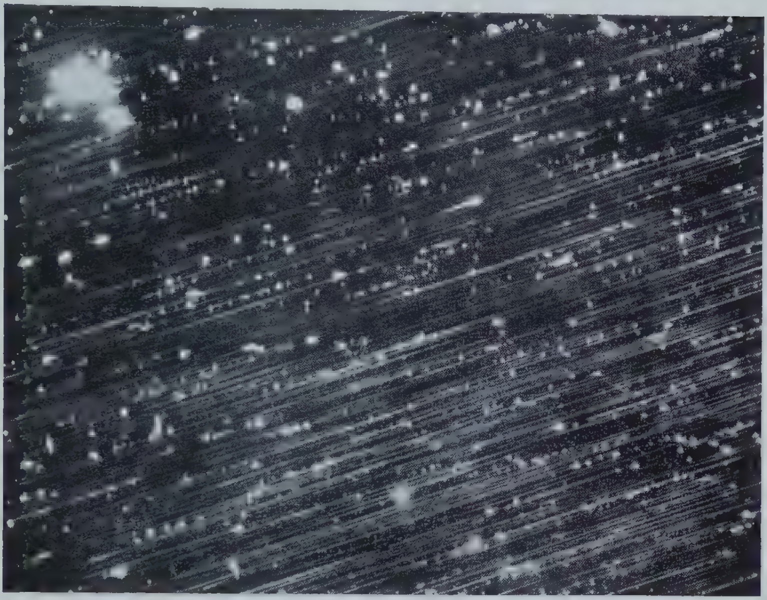


Fig. 1.—Appearance of slip lines on specimen 10F predeformed at 78 °K to 0·5, polished, and given an additional strain of 0·2 at 78 °K. (×150.)

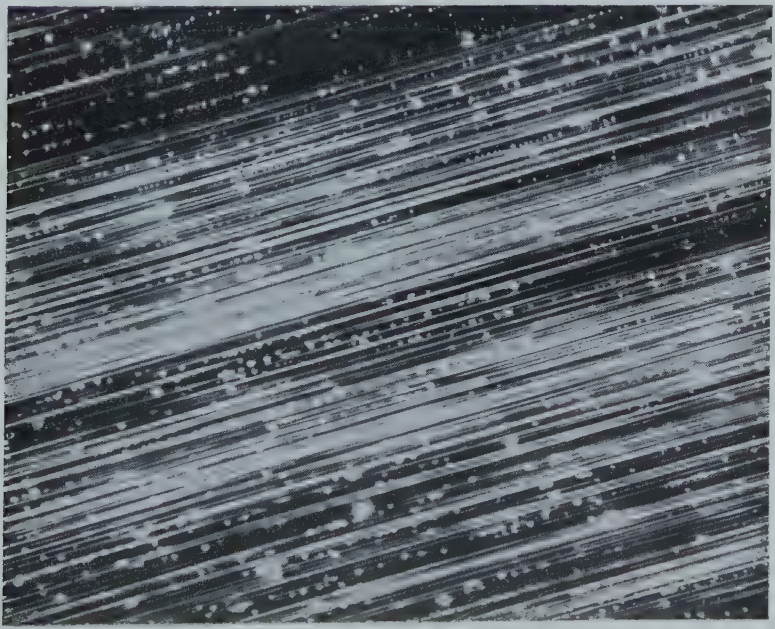
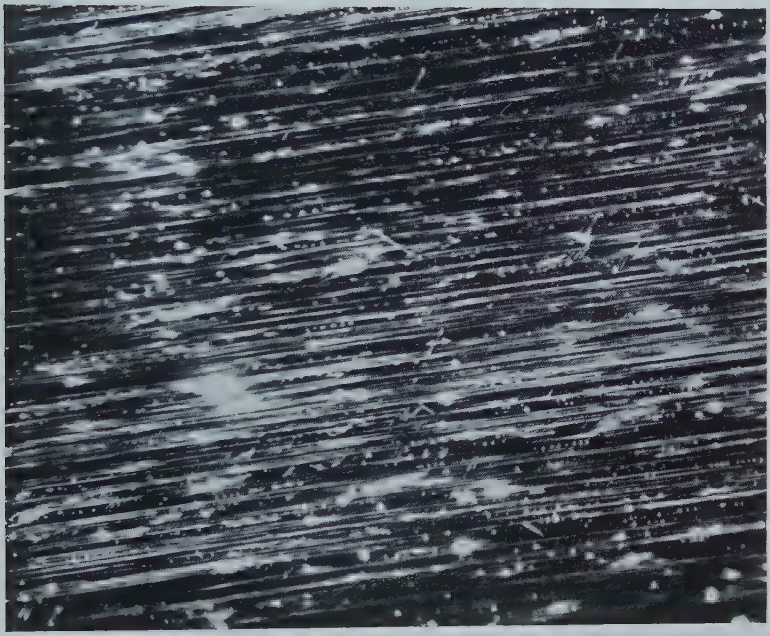


Fig. 2.—Slip in the work-softened specimen. (×150.)

FLOW STRESS OF MAGNESIUM SINGLE CRYSTALS



Slip lines formed in the rapid-hardening region at 78 °K. (×150.)

only at temperatures above about 100 °K. Above about 200 °K work softening was apparent at strains corresponding to stage B. In stage C work softening decreased somewhat. Work softening has previously been observed in hexagonal metals by Blaha and Langenecker (1958) in their work on zinc single crystals.

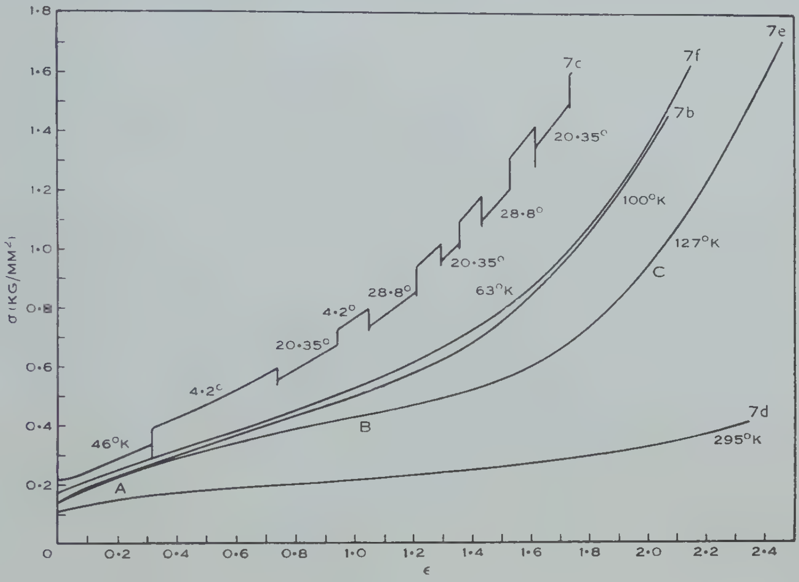


Fig. 6.—Stress-strain curves of specimens of orientation 7.

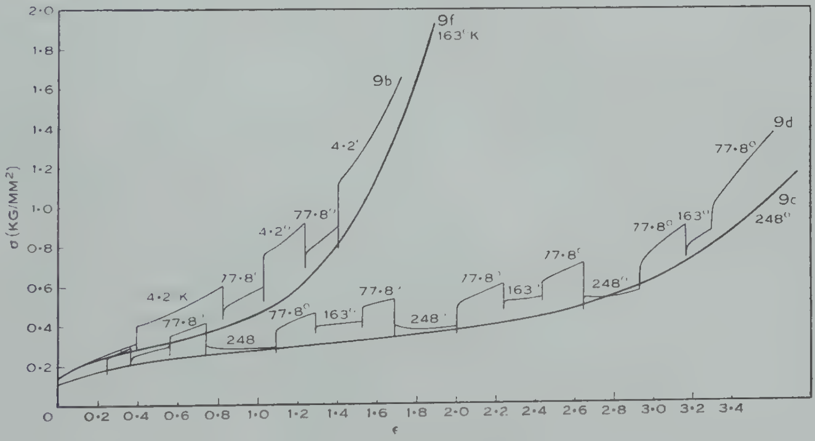


Fig. 7.—Stress-strain curves of specimens of orientation 9.

(*d*) *Surface Examination*

The surface of most specimens was examined microscopically after fracture. The specimens deformed at high temperatures usually showed long, prominent slip lines and quite a lot of deformation twins. Very often basal slip was observed in the twinned regions. At lower temperatures, where a higher rate of strain hardening was observed in stage C, the slip lines were usually shorter and fainter, comparable with the situation in face-centred cubic metals (e.g. Blewitt, Coltman, and Redman 1955), in this region the number of visible deformation twins decreased. Specimens deformed in liquid helium showed practically no deformation twins and the majority of slip lines had lengths of the order of 5×10^{-3} cm.

Some observations were made on the influence of strain and temperature on the appearance of the slip lines. At the beginning of deformation at room temperature, slip lines usually extended right across the specimen. At somewhat higher strains at liquid nitrogen temperature, they became fainter and shorter. Plate 1, Figure 1 shows the surface of specimen 10F which was predeformed at 78 °K to a strain of 0·5, chemically polished, and given an additional strain of about 0·2 at 78 °K. Plate 1, Figure 2 shows the surface of the same specimen after repolishing and straining by 0·2 at room temperature.

During this straining at room temperature the specimen work softened. It is clear from the figure that work softening in magnesium is not associated with cross slip, which seems to play such an important role in this phenomenon in face-centred cubic metals (e.g. Kelly 1956). The slip lines in this case seem to become coarser and bunch together.

At higher strains the slip lines become shorter, as illustrated in Plate 2, which shows the surface of the same specimen further deformed at 78 °K to a total strain of about 2, chemically polished, and strained at 78° by an additional strain of about 0·2.

IV. Discussion

No detailed theory of strain hardening in hexagonal metals has yet been proposed. It is usually ascribed, however, to large elastic interactions between parallel dislocations (e.g. Mott 1957 ; Seeger 1957). Conrad and Robertson (1957), in their work on magnesium single crystals, also reached this conclusion. Their measurements indicated that the flow stress changes were independent of, rather than proportional to, the total stress, in contrast to the present results which show a marked increase in $\Delta\sigma$ with increasing σ. This discrepancy could be due to the relatively small stress range covered in the former work, the variation of $\Delta\sigma$ lying within the experimental scatter.

The observed increase of $\Delta\sigma$ with σ and also the result that the corrected reversible flow stress changes by about 35 per cent. between 1·45 °K and about 250 °K show that obstacles acting over very short distances play a prominent part in the hardening process. Evidence against the elastic interaction theory has also been obtained by Edwards and Washburn (1954), who showed that hardening in the latent slip direction in zinc is higher than that in the active slip direction.

There is a very close resemblance between the form of the variation of $\Delta\sigma$ with σ in magnesium and face-centred cubic metals (Adams and Cottrell 1955 ; Cottrell and Stokes 1955 ; Basinski 1959). In both cases at higher stresses (stage γ for magnesium and stages II and III for face-centred cubic metals) $\Delta\sigma$ increases linearly with σ, although the range of stress values covered in magnesium is smaller since fracture occurs at comparatively low stresses. In the early stages of deformation in both magnesium and the face-centred cubic metals, the values of $\Delta\sigma$ are higher than expected, and, in magnesium, there seems to be another region of linear increase of $\Delta\sigma$ with σ, but having a somewhat steeper slope. The data at low temperatures in the face-centred cubic metals during the early stages of deformation are not yet sufficient to decide whether this second Cottrell-Stokes region also exists here. At very low temperatures in both cases, the Cottrell-Stokes law is obeyed over the entire deformation range.

This similarity in behaviour, coupled with a similar value of $\ln NkTAb\nu/\dot{\varepsilon}$, suggests that the fundamental obstacles and the thermally activated mechanism of overcoming them are very similar in hexagonal and face-centred cubic metals, and therefore, forest dislocations (Basinski 1959 ; Hirsch 1959) are likely to provide the main obstacles to slip also in the hexagonal metals.

The form of the $\Delta\sigma/\sigma$ curves, however, suggests that the random forest mechanism gives an oversimplified picture and that some other obstacles also influence the behaviour at low stresses and temperatures. A number of effects, e.g. cutting through dispersed vacancies produced by the deformation (Seitz 1952) or the difficulty of moving jogged dislocations (Maddin and Cottrell 1955), could be responsible. More data, especially on the influence of the thermal history of the specimen on the $\Delta\sigma/\sigma$ relation, are required before this question can be solved.

The similarity between the hexagonal and face-centred cubic metals is further illustrated by the similarity in their hardening processes. Stress-strain curves for cadmium (Boas and Schmid 1930) and magnesium (Schmid and Siebel 1931 ; and the present work) show not only a slow-hardening region with a rate of hardening $(d\sigma/d\varepsilon)/\mu$ about $2\cdot0\times10^{-4}$, which corresponds to stage I in face-centred cubic metals, but also a region of rapid hardening (cf. stage II in the face-centred cubic metals) ; here, however, the rate of hardening $(d\sigma/d\varepsilon)/\mu\sim1\cdot7\times10^{-3}$ is only about half that in the face-centred cubic metals. In cadmium, a bending over in the curves was apparent at high stresses (Boas and Schmid 1930), which may correspond to stage III of the face-centred cubic metals.

It is now generally accepted that stage I in the face-centred cubic metals ends with the onset of appreciable glide on secondary slip systems, caused either by the stress concentrations produced by dislocations piling up on the main glide plane, or rotation of the lattice, during deformation, to the symmetrical position where two slip systems are equally favoured. In hexagonal metals which we consider here, equivalent glide systems do not exist, and non-basal slip is more difficult than that on the basal plane, we could therefore expect more deformation before it is activated to any appreciable extent, and also a somewhat smaller rate of strain hardening when turbulent deformation is reached. It therefore appears that the hardening processes in these two types of metal are not likely to be fundamentally different.

From the foregoing discussion, it appears that the mechanism of strain hardening in face-centred cubic and hexagonal close-packed crystals (which deform primarily by basal slip) is fundamentally the same. Any mechanism of strain hardening must therefore rely on processes which are possible in both types of lattice. The difficulties of accounting for how the forest forms and is stabilized have not been solved, and the most promising attack on the problem may be the simultaneous study of both face-centred cubic and hexagonal metals.

V. ACKNOWLEDGMENTS

The author is grateful to Miss J. Spong, who determined the orientations of the single crystals and carried out many of the calculations, and also to Mr. W. Fisher, who prepared the single crystals.

VI. REFERENCES

ADAMS, M. A., and COTTRELL, A. H. (1955).—*Phil. Mag.* **46** : 1187.
BASINSKI, Z. S. (1959).—*Phil. Mag.* **4** : 393.
BLAHA, F., and LANGENECKER, B. (1958).—*Z. Metallk.* **49** : 357.
BLEWITT, T. H., COLTMAN, R. R., and REDMAN, J. K. (1955).—" Defects in Crystalline Solids." p. 367. (Phys. Soc. : London.)
BOAS, W., and SCHMID, E. (1930).—*Z. Phys.* **61** : 767.
CONRAD, H., and ROBERTSON, W. D. (1957).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **209** : 503.
COTTRELL, A. H., and STOKES, R. J. (1955).—*Proc. Roy. Soc.* A **233** : 17.
EDWARDS, E. H., and WASHBURN, J. (1954).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **200** : 1239.
GLEN, J. W. (1956).—*Phil. Mag.* **1** : 400.
HAASEN, P., and KELLY, A. (1957).—*Acta Met.* **5** : 192.
KELLY, A. (1956).—*Phil. Mag.* **1** : 835.
MADDIN, R., and COTTRELL, A. H. (1955).—*Phil. Mag.* **46** : 735.
MAKIN, M. J. (1958).—*Phil. Mag.* **3** : 287.
MOTT, N. F. (1956).—*Phil. Mag.* **1** : 568.
MOTT, N. F. (1957).—" Dislocations and Mechanical Properties of Crystals." p. 458. (Wiley : New York.)
SCHMID, E., and BOAS, W. (1950).—" Plasticity of Crystals : with Special Reference to Metals." (F. A. Hughes & Co. Ltd. : London.)
SCHMID, E., and SIEBEL, G. (1931).—*Z. Elektrochem.* **37** : 447.
SEEGER, A. (1957).—" Dislocations and Mechanical Properties of Crystals." p. 243. (Wiley : New York.)
SEEGER, A., DIEHL, J., MADER, S., and REBSTOCK, H. (1957).—*Phil. Mag.* **2** : 323.
SEITZ, F. (1952).—*Advanc. Phys.* **1** : 43.
SLUTSKY, L. J., and GARLAND, C. W. (1957).—*Phys. Rev.* **107** : 972.
WHITE, G. K. (1959).—" Experimental Techniques in Low Temperature Physics." (Oxford Univ. Press.)

THE INFLUENCE OF TEMPERATURE AND STRAIN RATE ON THE FLOW STRESS OF ANNEALED AND DECARBURIZED IRON AT SUBATMOSPHERIC TEMPERATURES

By Z. S. Basinski* and J. W. Christian†

[*Manuscript received September 25, 1959*]

Summary

The influence of temperature and strain rate on the flow stress of annealed and decarburized iron in the stages of uniform deformation has been studied at temperatures down to 78 °K. The Cottrell-Stokes law is not obeyed in iron, and the results are explainable in terms of the strong Peierls-Nabarro force in this metal.

I. Introduction

The temperature and strain-rate dependence of the yield stresses of body-centred cubic metals have been widely investigated. Comparatively little work, however, has been reported on the influence of these variables in the stages of uniform deformation corresponding to the movement of unpinned dislocations. Since the properties of free dislocations determine the deformation characteristics, and indirectly play an important part in such phenomena as brittle fracture, it would be desirable to separate the effects of temperature and strain rate on pinned and moving dislocations, and also to establish the extent to which interstitial impurities affect the properties of free dislocations.

There are two ways in which temperature can affect the flow stress. The size of obstacles which dislocations have to overcome may be temperature dependent through the variation of elastic constants (Cottrell and Stokes 1955) and Peierls-Nabarro force (Dietze 1952) or the obstacles can be overcome with the help of thermal fluctuations. In the latter case there should be a correlation between the effect of temperature and strain rate on the flow stress.

Body-centred cubic metals differ from the face-centred ones in having a large temperature dependence of the yield stress and small temperature dependence of the rate of strain hardening. Flow stress measurements in the face-centred cubic metals lead to the Cottrell-Stokes law, according to which the reversible change of flow stress produced by an increment in temperature or strain rate is proportional to the total flow stress. The experiments described in this work show that body-centred cubic metals exhibit quite different behaviour.

II. Experimental Method

Polycrystalline iron specimens of diameter 0·71 mm and length 5 cm were tested in a hard beam tensile testing machine (Basinski 1959) in the temperature range from about 80 °K to about 320 °K. The material used was " Ferrovac "

---

* Division of Pure Physics, National Research Council, Ottawa, Canada.

† Department of Metallurgy, University of Oxford, England.

vacuum-melted iron supplied by Vacuum Metals Co., Cambridge, Mass The chemical analysis was : C, 0·0027% ; O, 0·0107% ; N, 0·0004% ; Al, 0·001% ; Co, 0·003–0·006% ; Cu, 0·001% ; Mn, 0·005–0·007% ; Mo, 0·003% ; Ni, 0·009% ; Si, 0·007% ; Pb, 0·01% ; and Sn, 0·005%. Some specimens were cold drawn into wires and then annealed for 1 hr at 630 °C, which produced a mean grain diameter of about 4×10^{-3} cm. The other specimens were decarburized for about 18 hr at 750 °C in wet hydrogen when 1·25 mm in diameter, and then drawn to 0·71 mm diameter and reannealed for 2 hr at 650 °C *in vacuo*. This treatment produced specimens with a mean grain diameter of about 8×10^{-3} cm.

TABLE 1

CONSTANT TEMPERATURE BATHS

Bath	Temperature (°K)
Liquid nitrogen	78
Liquid oxygen	90
Acetone (m.p.)	180
Acetone-carbon dioxide	203
Calcium chloride-ice	229
Sodium chloride-ice	255
Ice-water	273
Controlled temperature bath	308
Controlled temperature bath	326

For the tensile tests, the specimens were immersed in the constant temperature baths listed in Table 1. The actual temperatures were checked by means of a gas thermometer. In most experiments the strain rate was changed between 10^{-4} and 10^{-3} sec^{-1} several times at constant temperature, ending with a strain rate of 10^{-4} sec^{-1}. The specimen was then unloaded slightly and after changing the temperature the deformation was resumed at a strain rate of 10^{-4} sec^{-1}. After 1–2 per cent. deformation at the new temperature, the strain-rate changes were resumed, followed by a temperature change to the original temperature. In some cases strain-rate changes were made between 10^{-4} and 10^{-5} sec^{-1}, $2\cdot5\times10^{-5}$ and $2\cdot5\times10^{-4}$ sec^{-1}, and also 5×10^{-5} and 5×10^{-4} sec^{-1}.

When testing annealed specimens, it was not possible to obtain an appreciable range of uniform deformation at lower temperatures unless the specimen was first deformed through the Lüders region at 0 °C. The magnitude of the Lüders deformation was about 5 per cent. tensile strain in these specimens. For this reason ice point was always used as a reference temperature, and the changes were made between 0 °C and various other temperatures.

III. EXPERIMENTAL RESULTS

According to the Cottrell-Stokes law, the reversible change in tensile flow stress, $\Delta\sigma$, between two fixed temperatures at constant strain rate, $\dot{\varepsilon}$, is given by $(\Delta\sigma/\sigma)_{\dot{\varepsilon}}=$constant. This relation, and the corresponding one for strain-rate changes, $(\Delta\sigma/\sigma)_T$=constant, are not even approximately true for polycrystalline iron. The results for both sets of specimens may be described more accurately by the relations,

$$\left.\begin{aligned}(\Delta\sigma)_{\dot{\varepsilon}} &\approx \text{constant},\\ (\Delta\sigma)_T &\approx \text{constant},\end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

but there are some complicating factors which should be discussed in more detail.

A major difficulty in the interpretation of the results for flow stress coefficients in face-centred cubic metals is the existence under some conditions of a small apparent yield drop on increasing the strain rate or decreasing the temperature.

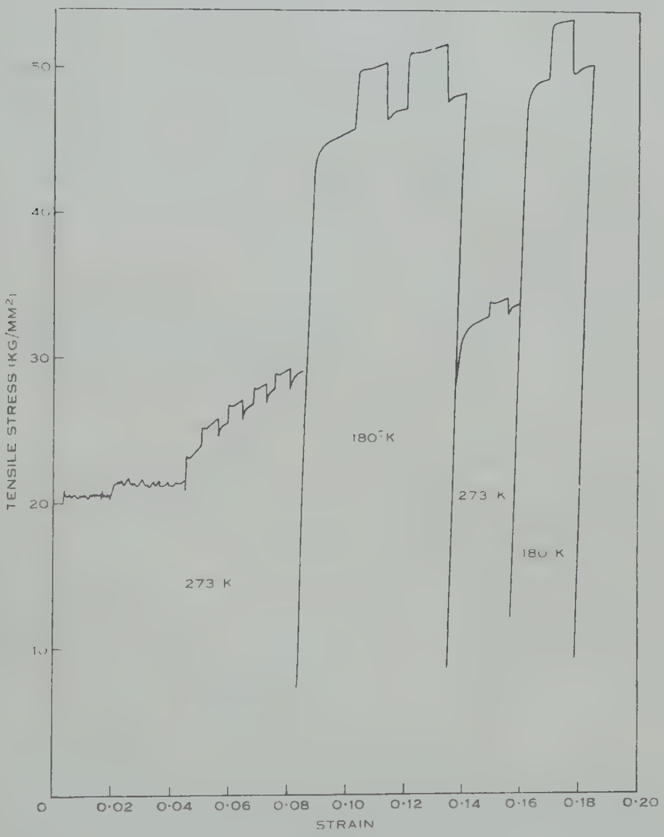


Fig. 1.—Typical stress-strain curve of annealed iron cycled between 273 and 180 °K.

In the present work this effect was serious only in annealed specimens at temperatures above about 200 °K. Figures 1 and 2 show typical portions of the stress-strain curves of annealed and decarburized specimens tested at 273 and 180 °K. In evaluating $\Delta\sigma$, upward changes in stress were always used, and σ values were taken from the immediately adjacent points on the strain axis.

Figure 3 shows plots of $\Delta\sigma$ against σ for decarburized specimens at 273, 180, and 78 °K. At 273 °K, $\Delta\sigma$ decreases approximately linearly over the whole range. In contrast to this, $\Delta\sigma$ shows a slight upward trend with increasing σ at 78 °K and is effectively constant at 180 °K. No systematic trend in $\Delta\sigma$ was ever observed for annealed specimens, but the available stress range in these tests is so small that the expected variation is smaller than the experimental scatter.

Figure 4 shows the $\Delta\sigma/\sigma$ plots at 273 °K for strain-rate changes of 10^{-5} to 10^{-4} sec^{-1}, $2{\cdot}5\times10^{-5}$ to $2{\cdot}5\times10^{-4}$, 5×10^{-5} to 5×10^{-4}, and 10^{-4} to 10^{-3}. The values of $\Delta\sigma$ produced by a tenfold increase in strain rate do not lie on a common curve but decrease with decreasing strain rate. The change in $\Delta\sigma$ is roughly proportional to the logarithm of the change in the lower strain rate. No variation in $\Delta\sigma$ was observed at 180 °K, and at 78 °K $\Delta\sigma$ seemed to increase slightly with decreasing strain rate, although the difference is about the limit of experimental error.

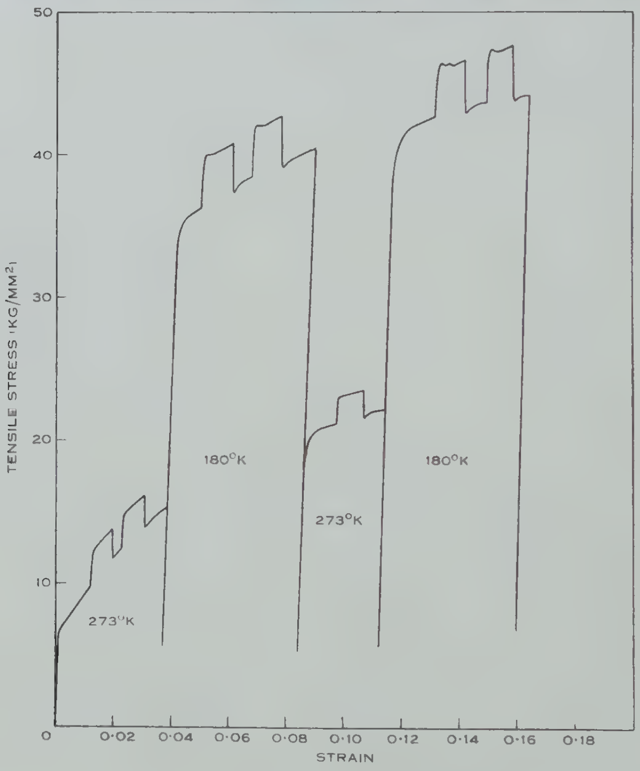


Fig. 2.—Typical stress-strain curve of decarburized iron cycled between 273 and 180 °K.

The variation in $\Delta\sigma$ with σ shown in Figure 3 and the influence of the actual strain rate on $\Delta\sigma$ may be correlated with the temperature variation of $(\Delta\sigma/\Delta \ln \dot{\varepsilon})$. This increases sharply with decreasing temperature, passes through a broad maximum in the range of 180 °K, and then decreases slightly again. We believe that the change in $\Delta\sigma$ with σ at constant temperature is a result of an increase in the number of moving dislocations contributing to the strain rate. At the beginning of deformation the number of available dislocations will be small and each dislocation will have to move at a relatively high speed in order to maintain the imposed strain rate. Multiplication of dislocations should enable them to move less rapidly, and this is equivalent to increasing the effective temperature.

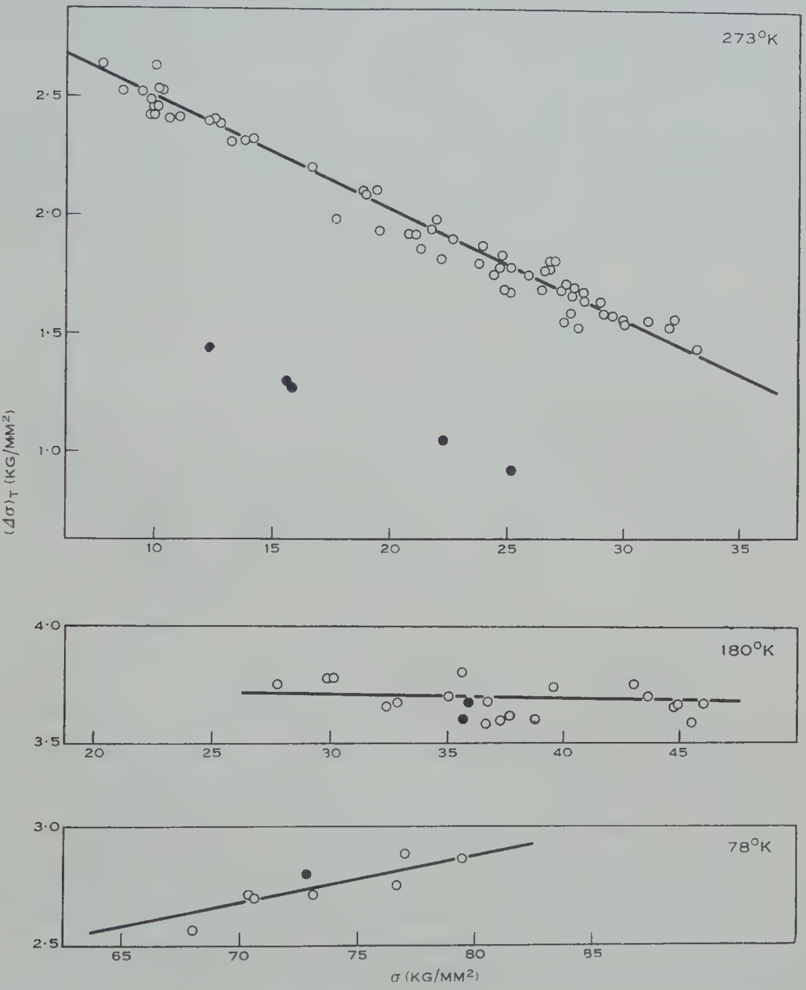


Fig. 3.—Plots, for decarburized iron, of the change of tensile stress produced by a strain-rate change of a factor of 10, versus total stress at 273, 180, and 78 °K. ○, represent strain-rate changes from 10^{-4} to 10^{-3} sec^{-1}; ●, represent strain-rate changes from 10^{-5} to 10^{-4} sec^{-1}.

This is in agreement with the sign of the change in $\Delta\sigma$ with strain rate. The proposed explanation would require about a tenfold increase in the number of active dislocations during the course of deformation.

The corresponding effect in face-centred cubic metals should be very difficult to observe, since the variation in $(\Delta\sigma)_T$ with temperature is small.

The way in which changes of flow stress produced by increments in the strain rate from 10^{-4} to 10^{-3} sec^{-1} are affected by temperature is shown in Figures 5 and 6, which are plots of $(1/T)\Delta\sigma/\Delta \ln \dot{\varepsilon}$ against T for annealed and decarburized specimens. In Figure 6 the values of $(1/T)\Delta\sigma/\Delta \ln \dot{\varepsilon}$ have been taken at 2, 10, and 20 per cent. strain. Within the experimental accuracy, at 20 per cent. deformation, $(1/T)\Delta\sigma/\Delta \ln \dot{\varepsilon}$ is a linear function of temperature. The results for the change in flow stress with temperature at a tensile strain rate of 10^{-4} are shown in Figure 7. The Cottrell-Stokes law was not obeyed, instead a result corresponding to equation (**1**) was obtained.

Some interesting observations were made incidentally in this work on the existence of the yield point drops in the stress-strain curves of decarburized specimens. This effect is usually attributed to the locking of dislocations by residual impurities (e.g. Paxton and Churchman 1953). The form of this yield drop at 180 °K is shown in Figure 8. The curve is very rounded and the typical Lüders deformation is completely absent. It is qualitatively very similar to the yield points produced by unloading and reloading the specimen. It is interesting to note that the yield drop is appreciable in the temperature range where $(\Delta\sigma)_T$ is large. These results suggest that the effect is a result of a rapid increase in the number of active dislocations at the beginning of deformation. The slope of the curve is consistent with a homogeneous process of this type rather than with the typical Lüders deformation expected from an impurity yield point.

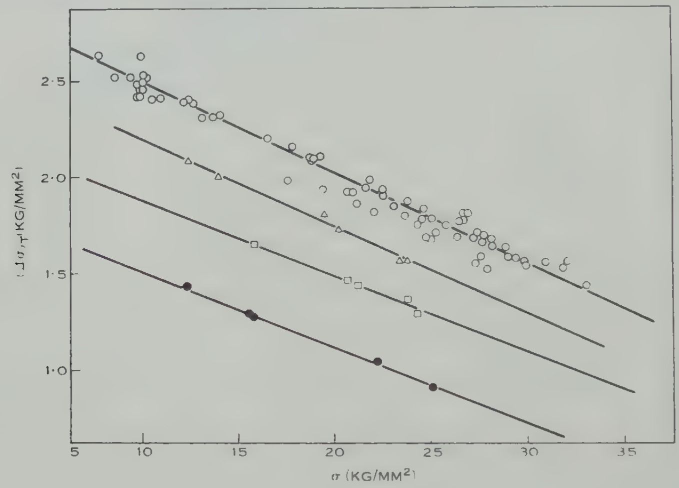


Fig. 4.—Plots of $(\Delta\sigma)_T$ versus σ. ○, for strain-rate changes of 10^{-4} to 10^{-3} sec^{-1}; △, for strain-rate changes of 5×10^{-5} to 5×10^{-4} sec^{-1}; □, for strain-rate changes of $2\cdot5\times10^{-5}$ to $2\cdot5\times10^{-4}$ sec^{-1}; ●, for strain-rate changes of 10^{-5} to 10^{-4} sec^{-1}.

IV. DISCUSSION

The form of the curves of $(\Delta\sigma)_T$ against σ show clearly that the obstacles determining the variation of flow stress with temperature and strain rate are not created during the deformation. Only two types of obstacle seem to satisfy this requirement, namely, a uniform distribution of impurities and the Peierls-Nabarro force of the lattice. Since the flow stress results for annealed and decarburized specimens are closely similar, despite the large differences in the initial yield, it is extremely improbable that interstitial impurity atoms are effective obstacles during deformation. Indeed, the concentration of interstitial impurities required to explain the magnitude of the observed flow stress at low temperatures is much too large. Assuming an interaction energy of 0·5 eV

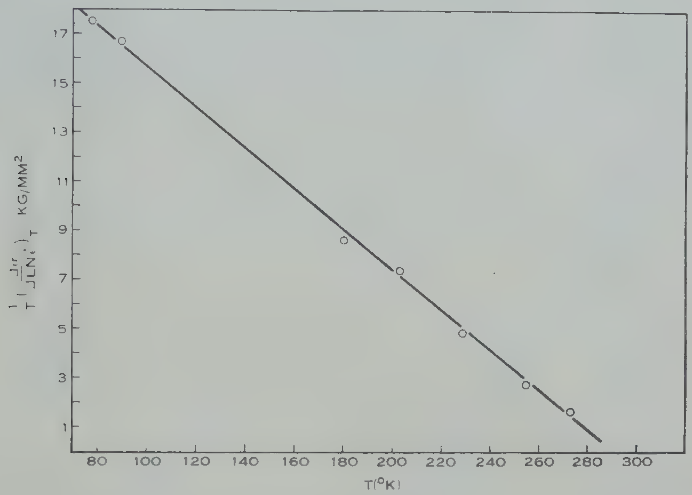


Fig. 5.—Temperature variation of $(1/T)(\Delta\sigma/\Delta \ln \dot{\varepsilon})_T$, for annealed iron.

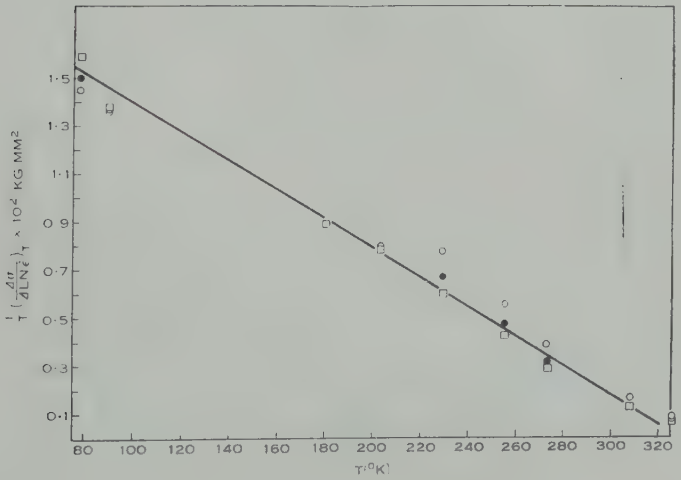


Fig. 6.—Temperature variation of $(1/T)(\Delta\sigma/\Delta \ln \dot{\varepsilon})_T$, for decarburized iron. ○, represent values taken at 2 per cent. strain ; ●, represent values taken at 10 per cent. strain ; □, represent values taken at 20 per cent. strain.

and an interaction distance of 2 Å between an interstitial and a dislocation, the separation of pinning points along a dislocation line is about 10 atoms, i.e. about 10 per cent. of impurity would be required.

These results seem to establish the importance of the Peierls-Nabarro force in iron. A similar conclusion was reached by Heslop and Petch (1956) from a study of the variation of the lower yield stress with grain size, impurity content, and temperature. The temperature variation ascribed by these workers to the Peierls-Nabarro force is different in detail from the present results, but the changes are of the same order of magnitude. Exact correspondence could not be expected since completely different techniques were used to obtain the relation between stress and temperature.

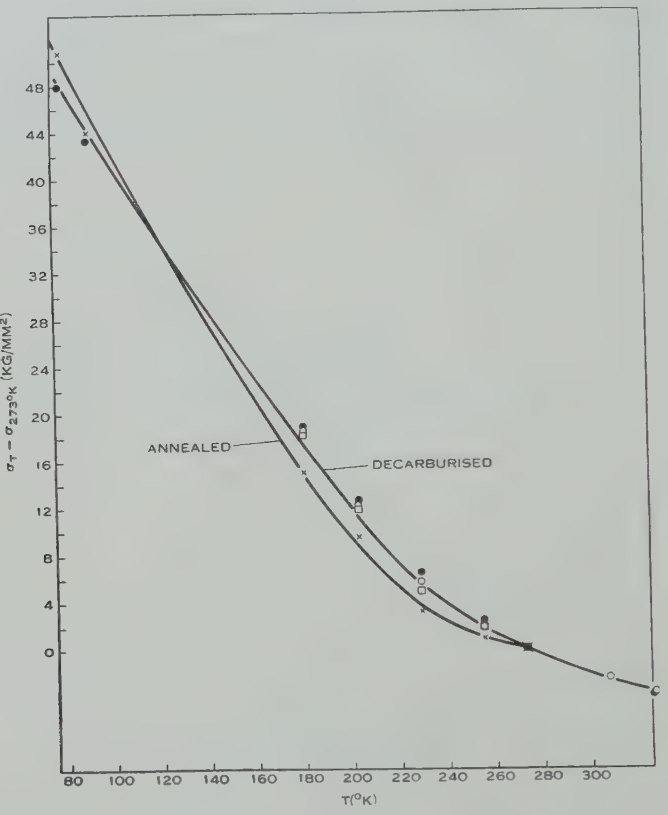


Fig. 7.—Temperature dependence of reversible flow stress in iron. ●, represent points taken at 2 per cent. strain; ○, represent points taken at 10 per cent. strain; □, represent points taken at 20 per cent. strain.

It is generally accepted that, provided the activation energy U is sensibly constant at all activation sites, the rate of thermally activated plastic flow may be written

$$\dot{\varepsilon}=NAb\nu e^{-U/kT}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

where N is the effective number of activated sites per unit volume, A is the area of the glide plane swept out by the dislocation after the activation process, b is the Burgers vector, ν is the characteristic frequency, and U is the energy of activation, which may be temperature dependent.

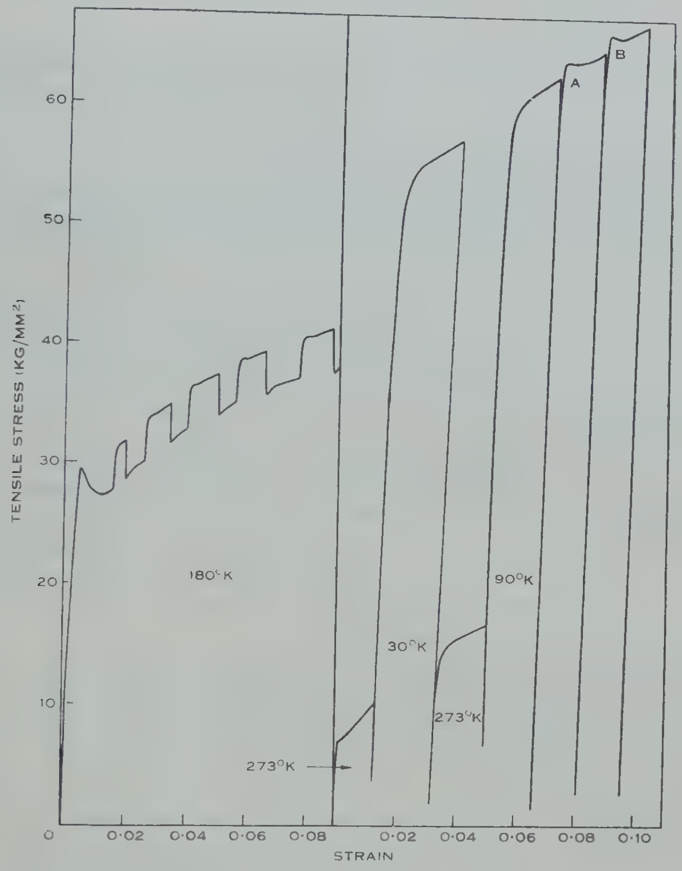


Fig. 8.—Appearance of yield drops in decarburized iron at various temperatures. A and B are unloading yield points at 90 °K.

In view of the large temperature variation of the flow stress in iron, and the fact that $(\Delta\sigma)_T$ is constant, it is permissible to neglect the influence of the temperature correction of the shear modulus on the variation of the flow stress with temperature. From equation (1) it follows that,

$$\ln (NAb\nu/\dot{\varepsilon})=(\partial\sigma/\partial T)_{\dot{\varepsilon}}\Big/\frac{1}{T}(\partial\sigma/\partial \ln \dot{\varepsilon})_T, \quad \ldots\ldots\ldots\ (3)$$

and

$$U=kT\ln (NAb\nu/\dot{\varepsilon}).$$

If the change of flow stress with temperature is due entirely to thermally activated processes, equation (3) gives a relation between $(\Delta\sigma)_{\dot{\varepsilon}}$ and $(\Delta\sigma)_T$.

The present results show that this is true for the deformation in iron. The temperature dependence of the flow stress agreed quite well with that expected from the strain-rate data assuming the validity of equation (**3**) and putting $\ln (NAb\nu/\dot{\varepsilon})=27$. The implicit assumption that this logarithm is constant is justified since none of the terms under the logarithm should be appreciably dependent on temperature.

The differences in the characteristics of thermally activated deformation in face-centred cubic and body-centred cubic metals is further illustrated in the value of $\ln (NAb\nu/\dot{\varepsilon})$ which is about 13–15 for the face-centred cubic metals (Basinski 1959) and 27 for iron. The larger value for iron is expected if we accept a model for the activation process similar to that introduced by Seeger (1956). Suppose the dislocation loops have polygonal form, and that the side of the polygon is of length L. Then L will be of the order of the slip line length and A will be approximately Lb. The activation process can take place at any part of a dislocation line and, if we assume that a length of line about 10 atoms long is involved in the activation, N becomes 5×10^6 per centimetre of dislocation. Assuming 10^7 dislocations/cm^3, this gives a value for the logarithm of about the observed magnitude.

The results show that the temperature variation of the flow stress can be accounted for in terms of thermally activated flow, and that the variation of the Peierls-Nabarro stress itself with temperature (Dietze 1952) is of comparatively minor importance.

V. REFERENCES

BASINSKI, Z. S. (1959).—*Phil. Mag.* **4**: 393.

COTTRELL, A. H., and STOKES, R. J. (1955).—*Proc. Roy. Soc.* A **233**: 17.

DIETZE, H. D. (1952).—*Z. Phys.* **132**: 107.

HESLOP, J., and PETCH, N. J. (1956).—*Phil. Mag.* **1**: 866.

PAXTON, H. W., and CHURCHMAN, A. T. (1953).—*Acta Met.* **1**: 473.

SEEGER, A. (1956).—*Phil. Mag.* **1**: 651.

THE PEIERLS–NABARRO FORCE AND THE DEFORMATION OF IRON AND STEEL

By N. Louat*

[*Manuscript received September 25, 1959*]

Summary

Evidence for the existence of an appreciable Peierls-Nabarro force in iron and steel is examined. It is concluded that the observed behaviour is unlikely to follow from the presently proposed mechanism of this force and that the observed frictional stress may arise as a consequence of dislocations being in pile-ups.

I. Introduction

Experimental evidence for the existence of an important Peierls-Nabarro force in iron arises mainly from the work of Petch and his collaborators on the variation of yield strength with grain size. Similar investigations of other transition group body-centred cubic metals are not very comprehensive but the indications are that these metals behave similarly. It is therefore probable that our present considerations of the behaviour of iron will also be pertinent to these other metals.

Petch (1953) showed that at a given strain rate the lower yield strength (σ) of iron and steel could be related to grain size by

$$\sigma=\sigma_0+kl^{-\frac{1}{2}}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where σ_0 and k are independent of the mean grain diameter but dependent on temperature. This relation has also been obtained theoretically by Petch (loc. cit.) and by Stroh (1955), who consider that deformation proceeds from one grain to the next when the shear stress at a certain position (S) ahead of a dislocation pile-up at a grain boundary reaches some characteristic value. This value, which enters the theoretically derived equation (**1**) through k, is supposed to be determined by the yield strength of a Cottrell-locked source at S and should therefore be sensitive to temperature and strain rate.

No experiments seem to have been made with varying strain rate but later experiments by Petch and others showed that k did in fact vary with temperature, although not in the manner expected. This work also showed that part of σ_0 could be attributed to various forms of solute hardening and was temperature independent, whilst the remainder (τ) was strikingly dependent on temperature. Heslop and Petch (1956) have further shown that it is improbable that the observed variation in τ could arise from any of the presently postulated dislocation models of work-hardening and, in the absence of any apparent alternative, these authors suggested that τ might arise from the existence of an appreciable Peierls-Nabarro force. This idea appears to have gained fairly general acceptance, although no detailed examination of the evidence seems to have been made. An examination is made in the present paper. We shall find that the temperature

---

* Aeronautical Research Laboratories, Department of Supply, Melbourne.

variation of σ_0 in iron is unlikely to be accounted for directly by the Peierls-Nabarro force and that this variation probably arises when the dislocations are in pile-ups and as consequence of Cottrell locking.

II. Examination of Experimental Results

Heslop and Petch (loc. cit.) measured yield strength (at a constant rate of specimen extension) as a function of grain size at a number of different temperatures. Results for particular temperatures were plotted against $l^{-\frac{1}{2}}$ as required by equation (**1**), thus revealing the temperature dependence of σ_0. A part of the temperature-independent portion of σ_0 was eliminated by observing the temperature-independent (linear) variation in σ_0 produced by varying the amount of carbon plus nitrogen in solution and extrapolating to zero.

Remembering that here specimen extension occurs as a result of Lüders band deformation it may be seen that a constant extension rate is consistent with the constancy of the theoretically time-dependent quantity k in equation (**1**) only if the amount of strain in the lower yield point extension is constant also. As we shall see later, this strain varies with the yield strength so that the experimental technique does not conform to the conditions of equation (**1**). However, this discrepancy is probably not very important since, as is well known, the yield strength is not very sensitive to the testing rate.

Although Heslop and Petch and others have found that their experimental results on the variation of yield strength with grain size are in reasonable agreement with an "$l^{-\frac{1}{2}}$" law, Baldwin (1958) has pointed out that the experimental data for a variety of materials including iron are equally well represented by straight lines when plotted against l^{-1}, $l^{-\frac{1}{2}}$, or $l^{-\frac{1}{3}}$. While this lack of definition can be ascribed to experimental scatter and to the small effective range of grain diameters available it is clear that the Petch-Stroh interpretation cannot be accepted unequivocally without further evidence. Some such evidence is to be found in the work of Lean, Plateau, and Crussard (1958) and that of Hutchison (personal communication, 1959). Lean, Plateau, and Crussard have shown for the case of a 0·07 per cent. carbon steel that the yield strength and yield elongation in a simple tension test are related through

$$\sigma-\sigma_i=D\varepsilon, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{2})$$

where σ_i and D are constants independent of test temperature and strain rate. This result may well be of general validity since an equation of this form may be deduced (Lean, Plateau, and Crussard loc. cit.) supposing, as in the Petch-Stroh model, that the strain results from the formation of dislocation pile-ups which terminate at the grain boundaries. If equations (**1**) and (**2**) are valid they should be compatible. Combining them and supposing that D is independent of grain size* and that $\sigma_0=\sigma_i$ we have

$$\varepsilon \propto l^{-\frac{1}{2}}. \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{3})$$

---

* This assumption implies that the number of slip planes per grain is constant. It is difficult to show conclusively that it is valid, since the behaviour of pile-ups on parallel planes has not yet been worked out. However, it is clear that this arrangement is singular in that the geometry is independent of grain size. As a consequence, the mutual interference of pile-ups is also invariant with varying grain size; a conclusion implicit in the derivation of equation (**1**).

Tentative confirmation of this proportionality has in fact been obtained by Hutchison (personal communication, 1959). This agreement indicates not only that the Petch-Stroh model is valid but also that

$$\sigma_i = \sigma_0. \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

A difficulty is apparent here since only σ_0 is temperature dependent. However, it may be observed that if the equation governing Lean's results were in fact

$$\sigma - \tau_i - \bar{\sigma}_i = \bar{D}\varepsilon,$$

and if τ_i and σ had the same or at least similar temperature and strain-rate dependences the experimentally observed relation would at least approximate to $\sigma - \sigma_i = D\varepsilon$ as required, where $\sigma_i = \bar{\sigma}_i \sigma/(\sigma - \tau_i)$ and $D = \bar{D}\sigma/(\sigma - \tau_i)$. The difficulty inherent in equation (4) may thus be resolved if

$$\tau_i(T) \equiv \tau(T),$$

in which case $\sigma(T)$ and $\tau(T)$ should be similar.

Heslop and Petch do not report the variation of σ with temperature in sufficient detail to permit a direct comparison of $\tau(T)$ and $\sigma(T)$ as determined by them. Nevertheless, a similarity between $\tau(T)$ and $\sigma(T)$ may be deduced from a comparison of Figure 1 with Figures 2, 3, and 4. The results shown in Figure 1 are those of Heslop and Petch, whilst the others are representative of the variations in yield strength with varying temperature obtained with different irons and steels. The results of Figures 2, 3, and 4 are clearly similar and in fact are well represented (Louat 1956) by equations of the same algebraic form. The similarity apparent between $\tau(T)$ and the various σ, $1/T$ relations should therefore be relevant in indicating the existence of the predicted similarity. Again, referring to equation (1) it is clear that if $\sigma(T)$ and $\tau(T)$ are to be similar they ought also be similar to $k(T)$. The experimental evidence on this point is conflicting ; Heslop and Petch indicate that $k(T)$ is sensitive to temperature only at temperatures below about ~110 °K. Cottrell (personal communication), on the other hand, has observed a more uniform variation.

It is clear that more experimental evidence is required in order to establish the prediction that $\sigma(T)$, $\tau(T)$, and $k(T)$ are similar, but the consistency of the various experimental results in terms of the Petch-Stroh model, subject to the existence of such a similarity, suggests that in fact they are similar.

Further experiments are also necessary to determine whether or not τ depends on composition (exclusive of carbon and nitrogen in solution). The experimental evidence available suggests that τ must vary with composition : if τ were invariant a decrease in slope ought to be but is not apparent in the σ, $1/T$ relations (see Figs. 2, 3, and 4) at the temperature at which $\tau(T)$ itself flattens, i.e. ~110 °K. If then τ is to vary with composition so that Heslop and Petch's result is but one of a family it would appear necessary that the equation describing the temperature dependence should contain at least four parameters.

Difficulties in directly identifying τ as the Peierls-Nabarro force now appear. It is difficult to see (*a*) why the parameters determining the Peierls-Nabarro force would vary with composition and (*b*) how as many as four parameters could be involved. One of these parameters would be eliminated if the low temperature flattening in $\sigma(T)$ and $\tau(T)$ were to arise from the intervention of twinning as a new process of deformation. The low temperature value of τ would then be characteristic of the twinning process and its contribution to the yield strength would, in view of its insensitivity to temperature, be essentially constant, independent of the temperature at which twinning begins. In this case $\sigma(T)$ ought to exhibit a discontinuity at the onset of twinning and not just the flattening observed unless, improbably, another temperature independent resistance to twinning but not slip is available.

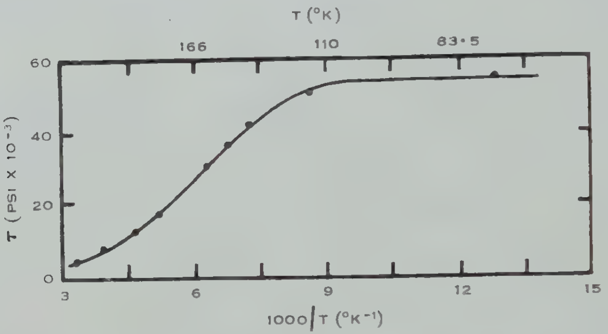


Fig. 1.—τ plotted against $1/T$ (Heslop and Petch 1956).

These considerations indicate that τ cannot be directly identified with the Peierls-Nabarro force as it is presently conceived.

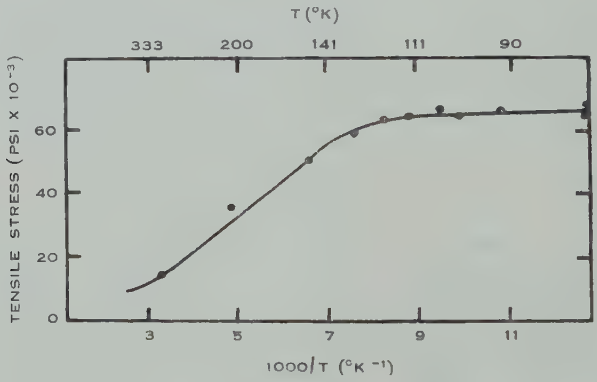


Fig. 2.—Yield strength plotted against $1/T$ (purified iron, Gibbons 1953).

It is clear that further experiments employing irons and steels of varying composition are needed to establish directly whether or not $\sigma(T)$, $\tau(T)$, and $k(T)$ are similar and whether or not τ varies with composition. Pending such investigations it is interesting to consider other possible processes which could lead to the observed results. In so doing we shall concentrate our attention on the

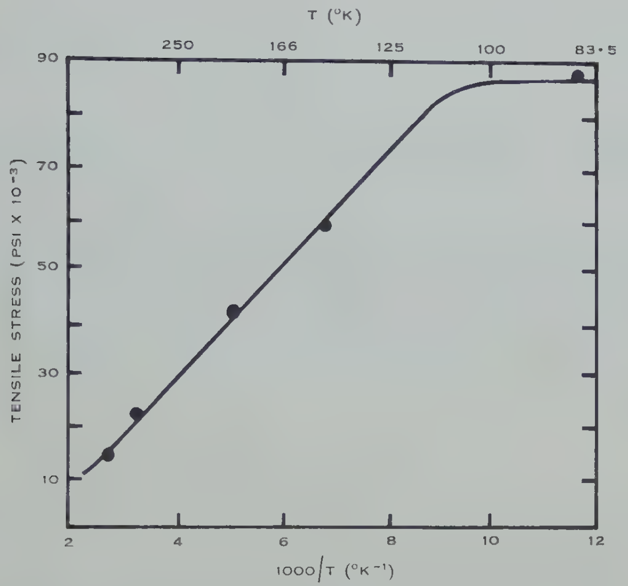


Fig. 3.—Yield strength plotted against $1/T$ (ingot iron, McAdam and Mebs 1943).

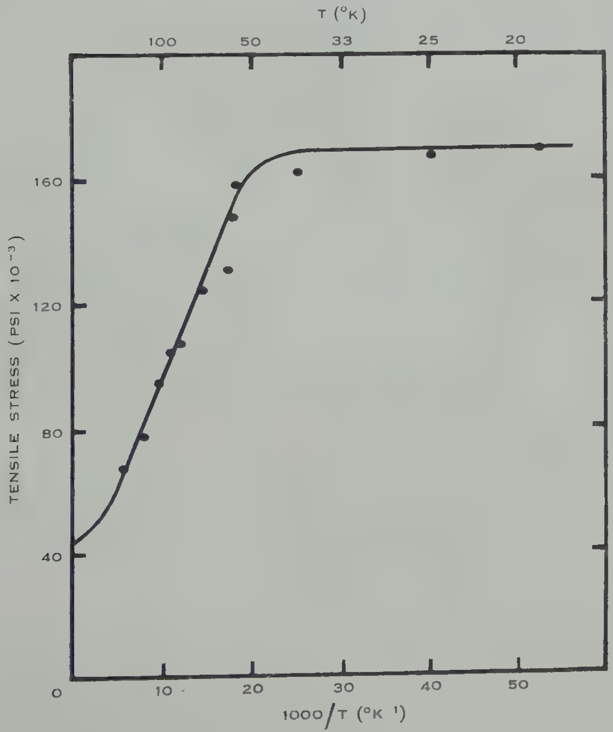


Fig. 4.—Yield strength plotted against $1/T$ (SAE 1020 steel, Eldin and Collins 1951).

anticipated similarity of $\tau(T)$ and $k(T)$ and look in the first instance for processes involving dislocation pile-ups and Cottrell locking as required by the Petch-Stroh model.

III. EFFECT OF PILE-UPS

Dislocations in a pile-up occupy equilibrium positions at which the shear stress is zero. The stresses near the dislocations are so arranged as to narrow them and hence increase the force necessary for motion.

As previously stated, an essential feature of the model used by Stroh and by Petch is that the secondary source is located near the head of the pile-up and experiences a stress which arises mainly from a comparatively few neighbouring dislocations. Eshelby, Frank, and Nabarro (1951) have shown that the number of dislocations between 0 (head of pile up) and x is given quite accurately, provided n/N is small, by

$$n=a\sqrt{(N\sigma_1 x)},\ \ \sigma_1=\sigma-\sigma_0,$$

where a is a constant and N is the total number of dislocations in the array. Subject to the condition $N\sigma_1$=constant, which is implied in the derivation of equation (**1**), it follows that the positions of the dislocations near the head of the array will be independent of the length of the pile-up and thus of the grain size. Similarly the (dislocation-narrowing) stresses near the nth dislocation of an array will vary only with $N\sigma_1$ and hence with temperature and strain rate as determined by the solute locking of the secondary source.

At present it does not seem possible to determine the form of the relation between a resultant τ and $N\sigma_1$ or even to make a reasonable estimate of its magnitude. A highly intuitive line of thought suggests a value of τ at the nth dislocation which varies like $N^2\sigma_1^2/n^2$ and which is equal or greater than the observed values of τ over a significant fraction of the length of the pile-up.

Again, it has been argued (Vreeland, Wood, and Clark 1953) that a solute-locked source must be activated each time it emits a dislocation. On this basis it is possible to derive a relationship between yield strength and grain size which may be a sufficient approximation to that given by Petch and by Stroh and which has the advantage that the frictional term is directly related to the yield strength. However, it is difficult to understand why the energy released when the dislocation rejoins its atmosphere after the first operation is not sufficient to ensure its continued movement.

As another process we note that deformation must involve more than one pile-up per grain on parallel planes. The behaviour of such arrays has not yet been worked out, but it is clear that the pile-ups will interfere with one another and introduce an apparent frictional stress. However, it is hard to see why this stress should be independent of grain size.

Finally, it is possible that an $\alpha\rightarrow\gamma$ phase change in iron, involving a decrease in volume of $\sim$6 per cent., could occur near the centre of a dislocation. Subsequent dislocation motion would involve an energy release and hence a frictional stress. The probability of this phase change would be only slightly greater when the dislocations are in pile-ups and for this reason the effect could be unimportant.

IV. CONCLUSIONS

It is concluded that τ is unlikely to arise directly from the Peierls-Nabarro model and it is suggested that the frictional stress is observed as a consequence of the yielding process. Consideration of a number of models to account for this indicates that the stress-induced narrowing of dislocations in a pile-up which modifies the Peierls-Nabarro force is the most plausible of them.

V. ACKNOWLEDGMENTS

The author wishes to thank the Chief Scientist, Australian Defence Scientific Service, Department of Supply, for permission to publish this paper.

VI. REFERENCES

BALDWIN, W. M., JR. (1958).—*Acta Met.* **6**: 139.
ELDIN, A. S., and COLLINS, S. C. (1951).—*J. Appl. Phys.* **22**: 1296.
ESHELBY, J. D., FRANK, F. C., and NABARRO, F. R. N. (1951).—*Phil. Mag.* **42**: 351.
GIBBONS, D. F. (1953).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **197**: 1245.
HESLOP, J., and PETCH, N. J. (1956).—*Phil. Mag.* **1**: 866.
LEAN, J. B., PLATEAU, J., and CRUSSARD, C. (1958).—*C.R. Acad. Sci., Paris* **247**: 1458.
LOUAT, N. (1956).—*Proc. Phys. Soc. Lond.* B **69**: 459.
MCADAM, D. J., and MEBS, R. W. (1943).—*Proc. Amer. Soc. Test. Mater.* **43**: 661.
PETCH, N. J. (1953).—*J. Iron Steel Inst.* **173**: 25.
STROH, A. N. (1955).—*Phil. Mag.* **46**: 198.
VREELAND, T., JR., WOOD, D. S., and CLARK, D. S. (1953).—*Acta Met.* **1**: 414.

THE ORIENTATION DEPENDENCE OF WORK-HARDENING IN CRYSTALS OF FACE-CENTRED CUBIC METALS

By L. M. Clarebrough* and M. E. Hargreaves*

[*Manuscript received September* 25, 1959]

Summary

It is shown that the principal features of the observed orientation dependence of work-hardening can be accounted for in terms of the likelihood of formation of Lomer-Cottrell sessile dislocations in two directions in the primary slip plane. This is deduced from the known variation of resolved shear stress with orientation, for the possible secondary slip systems, and metallographic observations of slip and deformation bands.

I. Introduction

The theories of work-hardening of Friedel (1957) and of Seeger *et al.* (1957), for crystals of face-centred cubic metals, propose that the interaction of piled-up groups of dislocations is the principal mechanism of hardening in the second stage of tensile deformation, represented by a relatively steep, linear stress-strain curve. The obstacles responsible for the piled-up groups are held to be Lomer-Cottrell sessile dislocations. Friedel has suggested that to effectively block activity in any slip plane it is necessary to form sessile dislocations in at least two directions in the slip plane ; the active sources can then be surrounded by these barriers.

In the present paper the consequences of these suggestions with respect to the orientation dependence of work-hardening will be examined. A brief statement of the experimental situation is given first. This is followed by a detailed examination of the likelihood of forming Lomer-Cottrell sessile dislocations in the primary slip plane, considering the known variation with orientation of the relative resolved shear stress for the various secondary systems and also the metallographic evidence available. It will become apparent that for all orientations the likelihood of forming such sessile dislocations in *one* direction in the primary slip plane is high. It is suggested that the variation in hardening with orientation therefore attaches to the likelihood of forming sessile dislocations in a *second* direction. The use of this criterion is shown to give qualitative agreement with the experimental observations.

II. The Experimental Situation

It is now generally accepted that the stress-strain curve of a face-centred cubic crystal in tension shows three stages in the most general case, although any of the stages may be absent in particular experiments because of the influence of variables such as orientation, temperature, and purity. See, for example, Masing and Raffelsieper (1950), Lücke and Lange (1952), Rossi (1954), Blewitt,

---

* Division of Tribophysics, C.S.I.R.O., Melbourne.

Coltman, and Redman (1955), Diehl (1956), Suzuki, Ikeda, and Takeuchi (1956), and Andrade and Aboav (1957).

An idealized curve for the general case is shown in Figure 1. The manner in which this curve varies as the orientation of the crystal is altered may be briefly summarized as follows.

In stage I, the easy glide region, the rate of hardening is low for crystals whose tensile axes are close to, but not at, the [110] corner of the standard stereographic triangle and increases as the orientation departs from this orientation towards the [100] or [111] corners or the [100]–[111] boundary. The observed rates of hardening vary by an order of magnitude; for orientations near the [100] and [111] corners the rate of hardening in stage I may be approximately half that in stage II and in these cases stage I is short and difficult to distinguish.

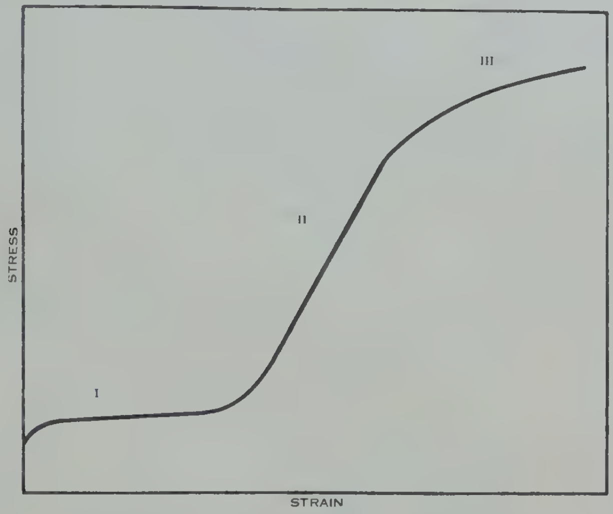


Fig. 1.—Schematic stress-strain curve for face-centred cubic crystals in tension.

For stage II the dependence of the rate of hardening on orientation is the same but the rate varies over all orientations only by a factor of two. For orientations which show a low rate of hardening in stage I the rate increases at least tenfold when stage II begins, in contrast to those orientations near [100] and [111] referred to above where it is merely doubled. The behaviour is perhaps best illustrated in the results of Diehl (1956) for copper, as quoted by Seeger (1957), shown in Figure 2.

For stage III there is no marked dependence of the rate of hardening on orientation, except in one respect. Crystals whose axes lie at or near [111] harden much more rapidly in stage III than those of other orientations. For example, the results of Lange and Lücke (1953) for aluminium show how different is the behaviour of a crystal oriented near [111] from one near [100] (Fig. 3). This trend is also discernible in the results of Lücke and Lange (1952) and those of Diehl (1956), particularly in their extended form as quoted by Seeger (1957).

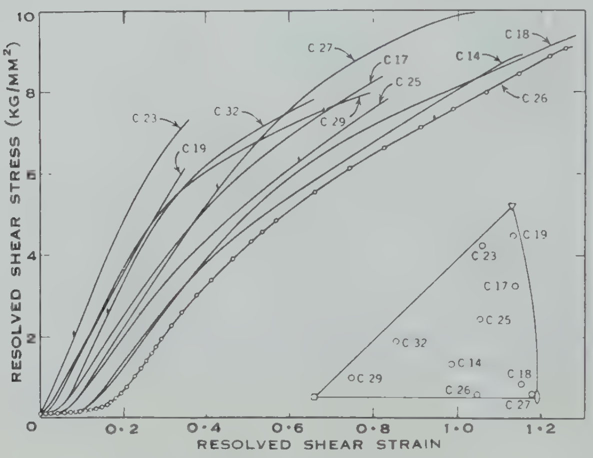


Fig. 2.—Stress-strain curves for crystals of copper of various orientations deformed in tension. Diehl (1956) as quoted by Seeger (1957).

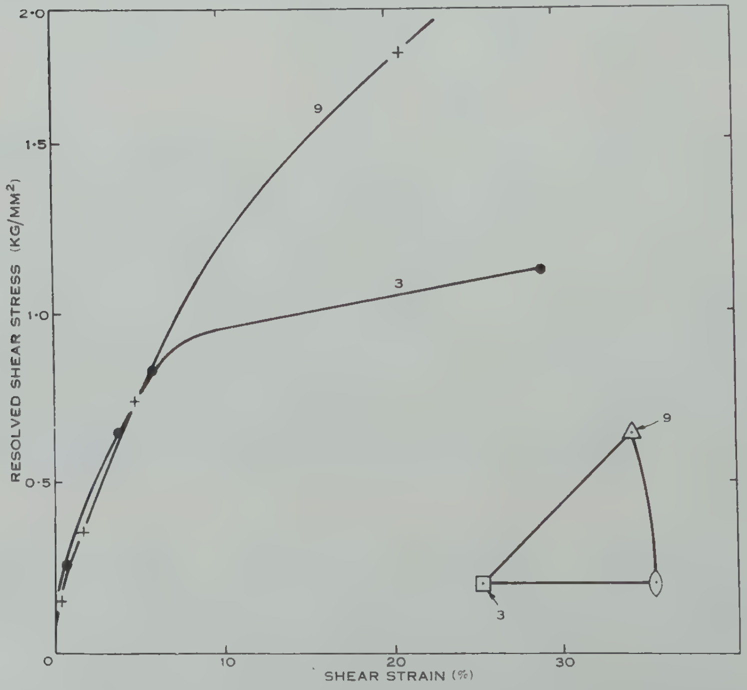


Fig. 3.—Stress-strain curves for crystals of aluminium of [100] and [111] orientations deformed in tension. Lange and Lücke (1953).

The principal features of the orientation dependence of work-hardening are therefore for stages I and II, the low hardening rate observed for orientations near, but not at, [110] and the increase in rate as the orientation changes towards the [100] or [111] corners or the [100]–[111] boundary of the triangle. For stage III the main feature is the continued rapid hardening of crystals oriented near [111].

III. The Probability of Forming Sessile Dislocations

In this section only sessile dislocations of the type considered by Lomer (1951) and Cottrell (1952) are discussed. Friedel (1955) has shown that these are the type most likely to occur because they are formed from pairs of dislocations which have both long-range and short-range attractive forces. The pairs of dislocations which can react to form such sessile dislocations in the primary plane are those whose Burgers vectors make an angle of 120° with each other, one member of the pair lying in the primary plane. These pairs of slip systems are listed in Table 1. It is understood that the pairs of systems with Burgers vectors in the same direction but opposite sense can also form sessile dislocations in the directions given in Table 1.

Table 1

COMBINATIONS OF SLIP SYSTEMS WHICH MAY FORM LOMER-COTTRELL SESSILE DISLOCATIONS IN THE PRIMARY PLANE

Direction of Sessile Dislocation Line	Pairs of Slip Systems
$[011]$	$(11\bar{1})[101]$–$(1\bar{1}1)[\bar{1}\bar{1}0]$ $(11\bar{1})[1\bar{1}0]$–$(1\bar{1}1)[\bar{1}01]$
$[1\bar{1}0]$	$(11\bar{1})[101]$–$(111)[01\bar{1}]$ $(11\bar{1})[011]$–$(111)[10\bar{1}]$
$[101]$	$(11\bar{1})[011]$–$(1\bar{1}\bar{1})[\bar{1}\bar{1}0]$ $(11\bar{1})[1\bar{1}0]$–$(1\bar{1}\bar{1})[01\bar{1}]$

Attention is concentrated on the primary plane because measurements of the change in orientation of a crystal during deformation indicate that, in general, slip on the primary plane makes the major contribution to the deformation; thus, it is in the primary plane that the formation of barriers to the movement of dislocations will have the greatest effect.

However, it is also apparent that even in the early stages of deformation secondary systems are active to some extent even though their presence may not always be detected by observation of the surface of the crystal. The more refined the technique of observation, the earlier are these secondary systems detected and evidence of this is presented by McKinnon (1960 and unpublished data). Also, there are the observations that precipitation from a supersaturated solid solution of silver in aluminium reveals the presence of secondary slip in the interior at very small strains, even at the beginning of easy glide (McKinnon 1955, unpublished data 1959).

These observations support the assumption made in the argument which follows that any secondary system will be active from the early stages of plastic deformation if the resolved shear stress on the system approaches that on the primary system. This activity need only be sufficient to produce enough dislocations on the secondary systems to create sessile dislocations. Use will be made of the curves of resolved shear stress for the various secondary systems, expressed as a fraction of the resolved shear stress on the primary system, calculated by Diehl *et al.* (1954) and reproduced in Figure 4. Ten secondary systems are included; for the eleventh, $(1\bar{1}\bar{1})[01\bar{1}]$, the relative resolved shear stress does not have appreciable values for any orientation and it was therefore omitted by Diehl *et al.*

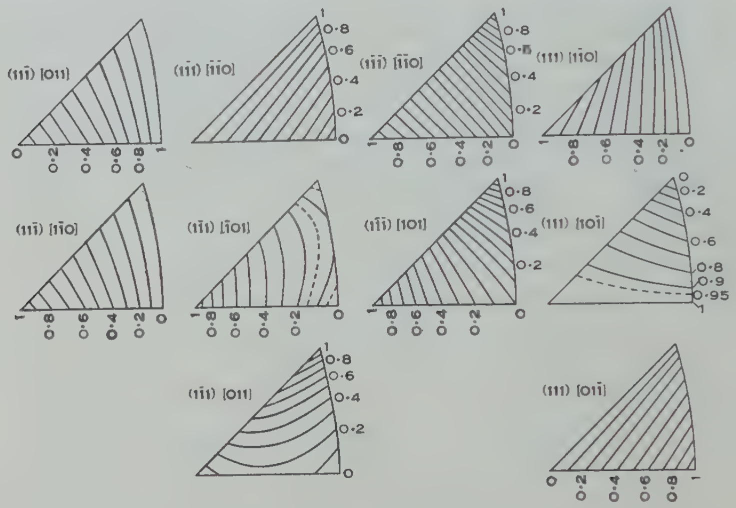


Fig. 4.—Contours of relative resolved shear stress as a function of orientation for ten secondary slip systems. The contours give the resolved stress on each secondary system as a fraction of that on the primary system for each orientation. After Diehl *et al.* (1954). The eleventh system, $(1\bar{1}\bar{1})[01\bar{1}]$, does not have appreciable relative resolved shear stress for any orientation.

We now consider which of the secondary systems that are capable of forming sessile dislocations in the primary plane are likely to operate for particular orientations, on the basis of the relative resolved shear stress on each system.

(*a*) *Orientations near [111]*

For crystals whose tensile axes lie near [111] rapid hardening is observed. From Table 1 and Figure 4 it can be seen that sessile dislocations will be readily formed in [011] directions by reaction of dislocations on the normal primary system $(11\bar{1})[101]$ and the conjugate system $(1\bar{1}1)[\bar{1}\bar{1}0]$, the latter having relative resolved shear stress of unity for the [111] orientation. The formation of sessile dislocations in the $[1\bar{1}0]$ direction is unlikely as both of the systems $(111)[01\bar{1}]$ and $(111)[10\bar{1}]$ have zero relative resolved shear stress for the [111] orientation. Sessile dislocations will be formed in the third direction [101] by the operation of the system $(11\bar{1})[011]$, a second direction in the primary plane, and the system $(1\bar{1}\bar{1})[\bar{1}\bar{1}0]$ in the cross plane, as both of these have high relative resolved shear stress at this orientation. Thus we expect sessile dislocations to be readily formed in [011] and [101] directions by the operation of systems on the primary, conjugate, and cross slip planes.

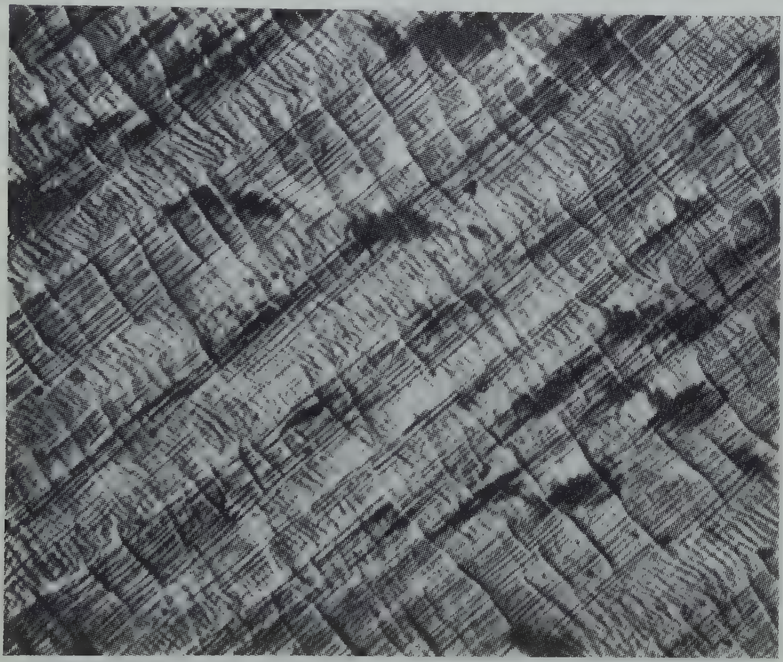


Fig. 5.—Slip markings on a crystal of aluminium oriented near [111] after 11·3 per cent. tensile strain. Sawkill and Honeycombe (1954).

The experimental situation is that these planes are observed to operate for this orientation, for example, by Rosi (1954). For orientations near [111] and close to the [111]–[110] boundary clustering of slip lines and bands of secondary slip are observed and the bands of secondary slip contain an unusually large proportion of traces of the primary plane. The behaviour is illustrated in Figure 5, due to Sawkill and Honeycombe (1954). This is attributed here to the operation of the system $(11\bar{1})[011]$ in the primary plane as predicted. The rotation produced by the operation of this system is such that it reduces the resolved shear stress on the normal primary system, and vice versa, so that once one system is established in a particular volume it tends to exclude the other, giving the observed bands of secondary slip. The experimental evidence is therefore consistent with the expected behaviour. It is observed that the rate of hardening of crystals decreases rapidly as the orientation departs from [111] along the [111]–[110] boundary. It will become apparent later that there is little opportunity for the formation of sessile dislocations in directions other than $[1\bar{1}0]$ as the orientation approaches [110]. The rapid decrease in the rate of hardening along the [111]–[110] boundary may be attributed to the fact that the relative resolved shear stress on the systems $(1\bar{1}1)[\bar{1}\bar{1}0]$ and $(1\bar{1}\bar{1})[\bar{1}\bar{1}0]$ necessary for the formation of sessile dislocations in the other two directions falls off rapidly along this boundary as the orientation departs from [111], Figure 4.

(*b*) *Orientations near [100]*

Crystals of these orientations are also observed to harden rapidly in stages I and II. Sessile dislocations in the [011] direction will be formed by reaction of dislocations on the normal primary and conjugate systems as in the previous case, but the systems $(11\bar{1})[1\bar{1}0]$ and $(1\bar{1}1)[\bar{1}01]$, which have high relative resolved shear stress for these orientations, may also contribute. However, a study of Figure 4 and Table 1 reveals that for any pair of slip systems capable of producing sessile dislocations in either of the other two directions one member of the pair has low relative resolved shear stress for these orientations.

This apparent anomaly is resolved by consideration of the metallographic evidence. It is clear from Figure 4 that the system $(111)[10\bar{1}]$ in the critical plane, has a very high relative resolved shear stress for a large range of orientations near the [100]–[110] boundary of the triangle. The operation of this system produces a rotation of the orientation which reduces the resolved shear stress on the primary system, and vice versa, and hence a tendency for mutual exclusion exists between these two systems. The experimental observation is that strong clustering of slip bands producing bands of secondary slip is always observed and even in macroscopic portions of the crystal the system to operate first may differ (Lange and Lücke 1953 ; Sawkill and Honeycombe 1954).

Thus, right from the beginning of deformation, there are bands where slip is mainly on the primary system alongside bands where slip on the system $(111)[10\bar{1}]$ in the critical plane predominates. This situation must give rise to considerable internal stresses, as the slip directions of the two systems make an angle of 90° with one another and the slip planes are inclined at 70°. These stresses will be in directions such that the systems $(111)[01\bar{1}]$ and $(11\bar{1})[011]$ will operate and sessile dislocations will be produced in $[1\bar{1}0]$ directions. Perhaps some will be produced in the [101] direction by reaction of the systems $(1\bar{1}\bar{1})[\bar{1}\bar{1}0]$–$(11\bar{1})[011]$, but except right at the [100] orientation the stress on the system $(1\bar{1}\bar{1})[\bar{1}\bar{1}0]$ is less than that on $(111)[10\bar{1}]$ so the sessile dislocations in $[1\bar{1}0]$ should predominate. In regions where the predominant system for primary slip is $(111)[10\bar{1}]$ the same considerations apply, having regard to the symmetry of the crystal, sessile dislocations being formed in $[1\bar{1}0]$ and $[01\bar{1}]$ directions in the critical plane.

It should be noted that the situation for orientations near [100] is quite different to that for those near [111], as in the latter case the main systems in the bands of primary and secondary slip lie in the same plane and their directions are inclined at 60°, so that high internal stresses are not produced.

The metallographic indications for crystals near [100] are in agreement with the above suggestions. A typical structure is shown in Figure 6 (Sawkill and Honeycombe 1954). Also, McKinnon (personal communication) has observed that the bands of secondary slip develop abnormally large tilts with respect to the surface for orientations along the [100]–[110] boundary, as would be expected if the internal stresses suggested here are relieved at the surface.

One might also expect that there would be exaggerated size-dependence of the rate of hardening for crystals near [100] if the dimensions of the crystal approached the scale of the inhomogeneities of deformation. Such size-dependence has, in fact, been found by Suzuki, Ikeda, and Takeuchi (1956) for crystals oriented near [100], the rate of hardening in stage II decreasing with the size, but this is reversed for crystals within 3° of [100]. For the exact [100] orientation the behaviour is very complex as eight slip systems are equally favoured and analysis is difficult. It is possible to see that sessile dislocations are readily created in directions lying in the (100) plane, but for others mechanisms of the type suggested above must be invoked.

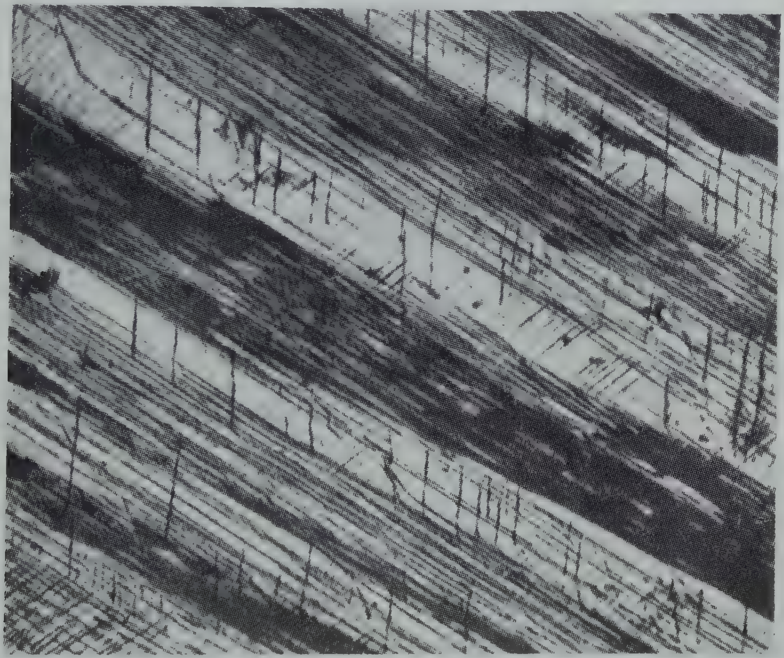


Fig. 6.—Slip markings on a crystal of aluminium oriented near [100] after 10 per cent. tensile strain. Sawkill and Honeycombe (1954).

(c) *Orientations along the [100]–[111] Boundary*

As the orientation departs from [100] towards [111] the inhomogeneities responsible for rapid hardening near [100] decrease in prominence, but at the same time the resolved shear stress on the system $(11\bar{1})[011]$ increases, while that on the system $(111)[10\bar{1}]$ remains high. Therefore, sessile dislocations can be formed in the $[1\bar{1}0]$ direction and the normal primary and conjugate systems will form sessile dislocations in the $[011]$ direction. As the orientation approaches $[111]$ the mechanisms characteristic of this orientation, which form sessile dislocations in the $[101]$ direction, will replace those which formerly formed sessile dislocations in the $[1\bar{1}0]$ direction, as the inhomogeneities change in character and the stress on the system $(111)[10\bar{1}]$ decreases. One expects, therefore, that there will be no great variation in rate of hardening along this boundary with perhaps somewhat lower rates near the mid point.

(*d*) *Orientations at [110]*

Crystals of these orientations have been studied rather less fully than those of other orientations, but it is known that crystals of the exact $[110]$ orientation harden rapidly. For this orientation four slip directions in two slip planes are equally favoured and sessile dislocations are readily formed in the $[1\bar{1}0]$ direction. Their production in a second direction would depend on inhomogeneities of deformation, as for orientations near $[100]$. Unfortunately, there is insufficient experimental evidence to allow detailed consideration of this case. It should be noted, however, that, apart from the systems $(11\bar{1})[011]$ and $(111)[10\bar{1}]$ which produce sessile dislocations in the $[1\bar{1}0]$ direction only, the system $(111)[01\bar{1}]$ has high relative resolved shear stress at this corner and this decreases rapidly as the orientation departs from $[110]$. The high rates of hardening observed at $[110]$ will not therefore persist as the orientation departs from this corner, as inhomogeneities of deformation of the type discussed above are not then observed.

(*e*) *Easy Glide Orientations*

These comprise the sector of the triangle beginning slightly away from $[110]$ and embracing the central portion of the triangle. This is approximately the area in which the conjugate system $(1\bar{1}1)[\bar{1}\bar{1}0]$ has low relative resolved shear stress and production of sessile dislocations in the $[011]$ direction by reaction with the primary system will not occur. The other pair of systems capable of forming sessile dislocations in this direction both have low relative resolved shear stresses in this area. We have seen that the systems $(11\bar{1})[011]$, $(111)[10\bar{1}]$ which are favoured in these orientations can produce sessile dislocations in the $[1\bar{1}0]$ direction, but the pairs of systems capable of forming sessile dislocations in the remaining $[101]$ direction each contain a member which has low relative resolved shear stress in this area. The most likely pair is $(11\bar{1})[011]$–$(1\bar{1}\bar{1})[\bar{1}\bar{1}0]$ and the second of these has zero resolved shear stress along the great circle joining $[110]$ and $[211]$, that is, through the centre of the sector being considered. There is thus little likelihood of producing sessile dislocations in a second direction for these orientations. The inhomogeneities of deformation observed are mainly kink bands, for which the internal stress can be accommodated by accumulation of dislocations on the primary system.

The operation of the primary system alters the orientation of the crystal towards the $[100]$–$[111]$ boundary and raises the stress on the conjugate system, thus increasing the efficiency of production of sessile dislocations in the necessary second direction. These considerations would predict that crystals at these orientations would deform initially with a low rate of hardening until sufficient rotation of the orientation occurred to bring the conjugate system into operation and cause an increase in rate of hardening. The further the initial orientation from the [100]–[111] boundary, the greater would be the strain before the increase in rate of hardening occurred. This is exactly the experimental result for easy glide. Further, Paterson (1955) has shown that, if deformation of the crystals is carried out in alternating tension and compression, which produces no overall rotation of the orientation, the cumulative strain which can be achieved at the low rate of hardening characteristic of easy glide is greatly increased.

(*f*) *Orientation Dependence in Stage III*

The experimental observation is that crystals whose orientations lie near [111] harden much more rapidly in stage III than all others; there is only a slight deviation from the linear curve observed for stage II. For other orientations a marked decrease in slope occurs. The difference between the [111] and [100] orientations is illustrated in Figure 3. The above consideration of the formation of sessile dislocations affords a ready explanation of this behaviour. It is only for orientations near [111] that the efficiency of production of sessile dislocations in the [101] direction is high. This direction is the intersection of the primary and cross planes and is the common slip direction for normal primary and cross slip. Extended sessile dislocations in this direction will therefore form obstacles to the cross slip of screw dislocations, which is suggested by Seeger *et al.* (1957) to be the mechanism by which the screw components of dislocations in piled-up groups circumvent the sessile dislocations and produce the decreased rates of hardening in stage III.

IV. DISCUSSION

The considerations outlined in the previous section afford a ready explanation of the main features of the orientation dependence of work-hardening in all stages of the stress-strain curve. It is suggested that the basic mechanism of hardening is the same in stages I and II. On this view, easy glide is terminated when sufficient sessile dislocations are formed to block off the active sources from the surface of the crystal and prevent dislocations leaving the crystal. For orientations which harden rapidly in stage II we have seen that the efficiency of production of sessile dislocations in the necessary two directions is high and stage I is therefore short and of high slope, as this blocking off is achieved at small strains. For orientations which show extensive easy glide, the efficiency of production of sessile dislocations in a second direction is low and it is not until rotation of the orientation of the crystal increases this efficiency that easy glide is terminated. Some dependence of the rate of hardening in stage II on orientation remains, but it is not nearly as great as in stage I, because of this effect of the change in orientation increasing the efficiency of production of sessile dislocations for orientations where this was formerly low.

For stage III the effect of sessile dislocations in the [101] direction as barriers to the cross slip of screw dislocations satisfactorily explains the higher rates of hardening observed for orientations near [111].

These conclusions have some relevance to other observed features of work-hardening. The fact that the rate of work-hardening in stage II is independent of temperature (Blewitt, Coltman, and Redman 1955) follows from the fact that the factors determining the rate are primarily geometrical and thus the rate depends only on the orientation.

The extent of easy glide for a given orientation is increased by lowering the temperature or by impurities in solid solution. The effect of temperature is explained following Seeger *et al.* (1957) by the suggestion that the dislocations on secondary systems must cut through dislocations on the primary system during the slip activity which produces sessile dislocations. This will be a temperature-dependent process, the ease of forming sessile dislocations decreasing with temperature and thus the extent of easy glide increases.

The effect of impurities also follows from the ideas proposed above. It has been suggested that this is related to Cottrell locking of secondary sources (Cottrell 1954; Rosi 1954; Smallman 1955). Once the yield stress is exceeded, the dislocations on the primary system are virtually unimpeded by the impurities, but the operation of secondary sources may be delayed by atmospheres of impurity atoms. This retards the production of sessile dislocations and results in an increase in the extent of easy glide.

V. REFERENCES

ANDRADE, E. N. DA C., and ABOAV, D. A. (1957).—*Proc. Roy. Soc.* A **240**: 304.

BLEWITT, T. H., COLTMAN, R. R., and REDMAN, J. K. (1955).—"Defects in Crystalline Solids." p. 369. (Phys. Soc.: London.)

COTTRELL, A. H. (1952).—*Phil. Mag.* **43**: 645.

COTTRELL, A. H. (1954).—"Relation of Properties to Microstructure." p. 131. (Amer. Soc. Metals: Cleveland, Ohio.)

DIEHL, J. (1956).—*Z. Metallk.* **47**: 331.

DIEHL, J., KRAUSE, M., OFFENHÄUSER, W., and STAUBWASSER, W. (1954).—*Z. Metallk.* **45**: 489.

FRIEDEL, J. (1955).—*Phil. Mag.* **46**: 1169.

FRIEDEL, J. (1957).—Discussion on work-hardening and fatigue in metals. *Proc. Roy. Soc.* A **242**: 147.

LANGE, H., and LÜCKE, K. (1953).—*Z. Metallk.* **44**: 183, 514.

LOMER, W. M. (1951).—*Phil. Mag.* **42**: 1327.

LÜCKE, K., and LANGE, H. (1952).—*Z. Metallk.* **43**: 55.

McKINNON, N. A. (1955).—*Phil. Mag.* **46**: 1150.

McKINNON, N. A. (1960).—*Nature* **186**: 148.

MASING, G., and RAFFELSIEPER, J. (1950).—*Z. Metallk.* **41**: 65.

PATERSON, M. S. (1955).—*Acta Met.* **3**: 491.

ROSI, F. D. (1954).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **200**: 1009.

SAWKILL, J., and HONEYCOMBE, R. W. K. (1954).—*Acta Met.* **2**: 854.

SEEGER, A. (1957).—"Dislocations and Mechanical Properties of Crystals." p. 243. (John Wiley and Sons: New York.)

SEEGER, A., DIEHL, J., MADER, S., and REBSTOCK, A. (1957).—*Phil. Mag.* **2**: 323.

SMALLMAN, R. E. (1955).—*J. Inst. Met.* **84**: 10.

SUZUKI, H., IKEDA, S., and TAKEUCHI, S. (1956).—*J. Phys. Soc. Japan* **11**: 382.

THE PLASTIC RESISTANCE OF CRYSTALS

By J. J. Gilman*

[*Manuscript received September 25, 1959*]

Summary

Evidence is presented in this paper to show that crystals present a *plastic resistance* to dislocation motion. This has been demonstrated by direct observations of dislocations in LiF crystals by Johnston and Gilman, and in Fe-Si crystals by Stein and Low. The plastic resistance is almost independent of dislocation velocity except at very high velocities. It is, however, strongly influenced by temperature, impurities, and radiation damage. In LiF crystals, the plastic resistance is much higher for glide on {100} planes than on {110} planes.

It is shown that the yield stresses and indentation hardnesses of many crystals can be correlated with their elastic moduli, provided the crystal structure type is held constant. For example, for face-centred cubic metals the yield stress is equal to $1 \cdot 6 \times 10^{-3} G$ (shear modulus) at —200 °C. This is taken to mean that one important factor that determines the plastic resistance is the elastic modulus of a crystal. Crystal structure type also has a big effect, so the detailed structure of the core of a dislocation also influences the plastic resistance.

The temperature dependence of the plastic resistance is discussed. A sharp drop in the plastic resistances of covalent crystals is observed at a temperature that is roughly equal to two-thirds of the melting temperature. It is suggested that this is because the cores of dislocations in covalent crystals " melt " at this temperature.

I. INTRODUCTION

One of the primary mechanical characteristics of a crystal is its yield stress. At this stress the behaviour of the crystal suddenly changes from being elastic to being plastic, which means that the plastic strain rate suddenly becomes finite and sometimes it becomes quite large. In terms of dislocation movements the plastic strain rate ($\dot{\gamma}$) depends on the Burgers vector (b) of the dislocations in the crystal, the average velocity of the dislocation motions ($\bar{v}$), and how many dislocations are moving per unit area (ρ). These quantities are related by a simple equation (Cottrell 1953)

$$\dot{\gamma} = b\rho\bar{v},$$

which has been verified by Johnston and Gilman (1959) through direct measurements of dislocation densities and velocities in LiF crystals. Since b is a constant and ρ cannot change unless dislocations move, the yield stress is related to $\bar{v}$. Thus, in order to understand the nature of the yield stress, the first question that requires an answer is: what determines the stress needed to move dislocations at a large average velocity? Is it the stress needed:

(*a*) to operate sources of dislocations such as Frank-Read sources?
(*b*) to pull dislocations away from Cottrell atmospheres?
(*c*) to drive dislocations through a forest of other dislocations?
(*d*) to overcome an intrinsic *plastic resistance* of the crystal structure?

---

* General Electric Research Laboratory, Schenectady, New York, U.S.A.

Also, do various types of crystals (metallic, ionic, or covalent) depend on different ones of these factors, or is plastic flow in all crystals dependent on the same factor? It is the thesis of this paper that factor (*d*) is the most important one for all single phase crystals. This thesis will be supported by experimental evidence.

A second question is: what are the most important properties of a crystal in determining its plastic resistance? It will be shown that the elastic modulus is a very important property, and lattice structure also plays an important part.

The final question that will be considered in this paper is: why does the plastic resistance of crystals decrease with increasing temperature?

II. QUESTION ONE: WHAT DETERMINES THE STRESS NEEDED TO MOVE DISLOCATIONS?

This question has been answered by direct observations of dislocations moving through crystals. In order to do this, techniques were developed by Gilman and Johnston (1957) for isolating individual dislocations in LiF crystals and observing their motions. The method uses etch-pitting for locating the dislocations before and after their movements. This is illustrated in Figure 1.

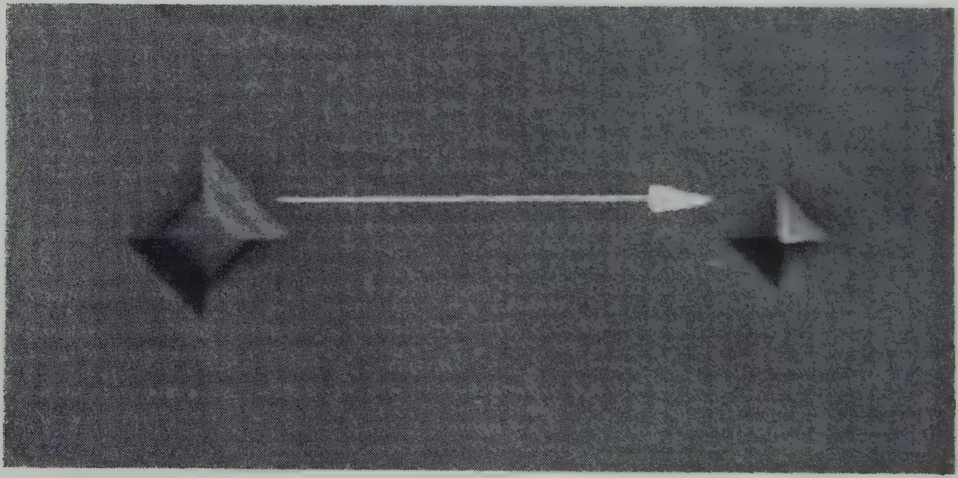


Fig. 1.—Movement of an isolated dislocation in a lithium fluoride crystal. The crystal surface was etched, then a stress was applied to move the dislocation, and finally the surface was re-etched. The initial dislocation position is shown by the flat pit; the final position by the pointed pit. 750×.

It has been found experimentally that dislocations accelerate very quickly and then move with constant velocity if a constant stress is applied to them (Johnston and Gilman 1959). Therefore, their motion can be summarized by velocity-stress curves. These curves have been determined for LiF crystals by applying stress pulses of various durations (from $\sim 10^{-6}$ to 10^{+5} sec) to isolated dislocations in the crystals. The crystals that contained the dislocations were etched before and after the stress pulses were applied so the distances that the dislocations moved could be measured. Average velocities of the dislocations at various stress levels were thus determined, and a plot of the results for a typical crystal is shown in Figure 2.

The behaviour shown in Figure 2 has several features. Two important ones are: (*a*) the fact that no dislocation motion occurs below a certain stress; (*b*) that, once having begun to move, dislocations increase their rates of motion very rapidly with further increases in the applied stress. Over the initial range of stresses the dislocation velocity increases in approximate proportion to the 25th power of the stress. (See Appendix I for further discussion of the stress dependence.) Another feature is the fact that edge dislocations move about 50 times faster than screw dislocations at the same stress. Then, since the time required for an edge dislocation to move a distance x is $t_e = x/v_e$ and the time for a screw dislocation to move the same distance is $t_s = x/v_s$, the average velocity with which an area x^2 can be swept out by a dislocation loop is $\bar{v} = 2x/(t_e + t_s) = 2v_e v_s/(v_e + v_s)$, but if $v_s \ll v_e$, then $\bar{v} \simeq 2v_s$. Therefore, it is the motion of the slower moving screw dislocations that controls the overall rate of flow.

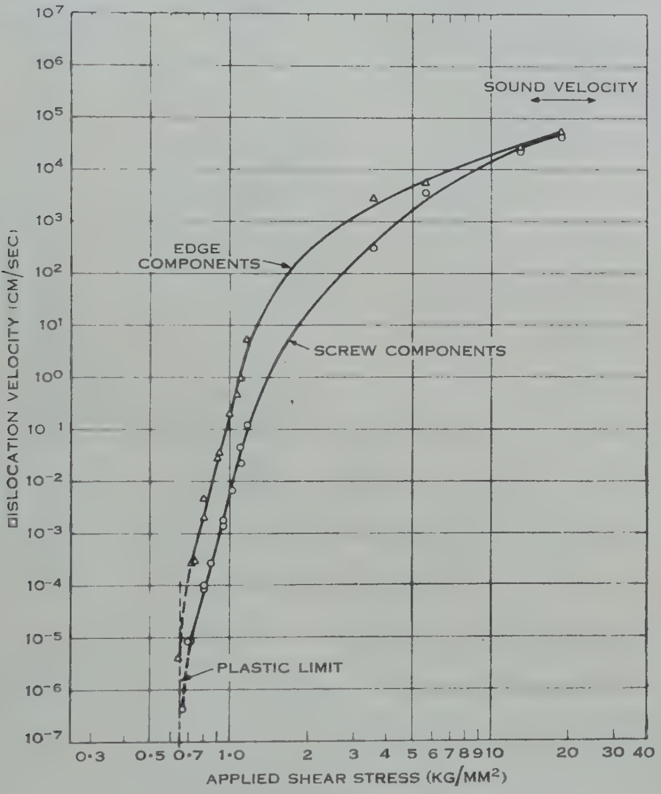
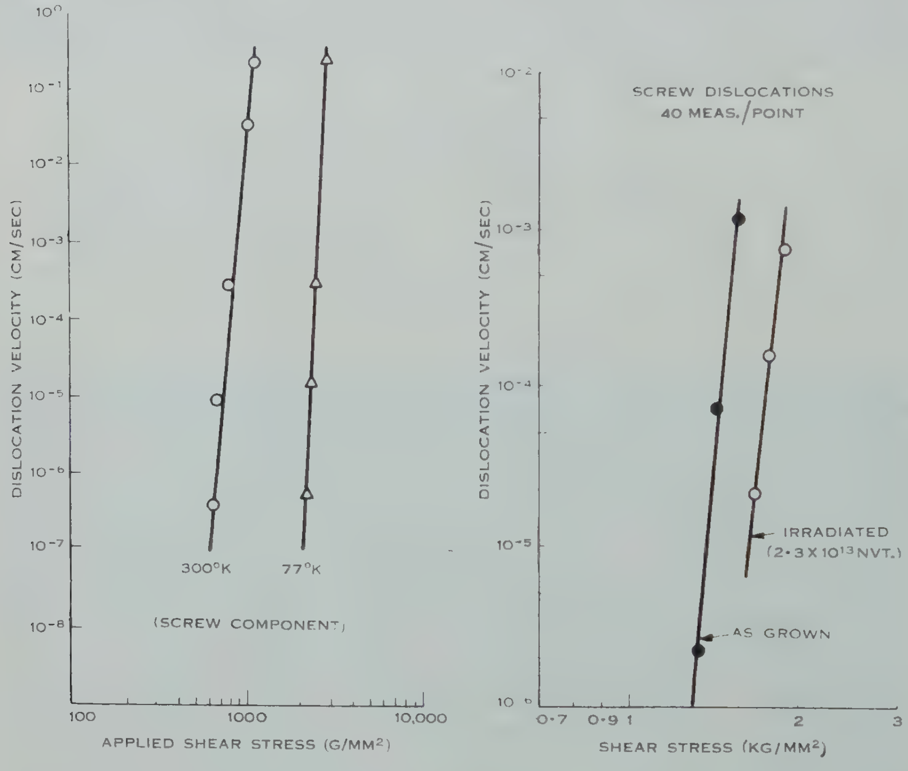


Fig. 2.—Dislocation velocities in a typical LiF crystal.

It may also be seen (Fig. 2) that dislocation motion is strongly damped in LiF crystals. If the applied shear stress is τ, then the energy that is dissipated per unit distance of motion is τb. This results because the work done on the dislocation must equal the energy dissipated if the velocity is constant. Therefore, there is a quasi-viscosity or *plastic resistance* associated with dislocation motion in LiF. It seems natural that the plastic resistance should determine the flow behaviour of LiF crystals, but in order to prove this point unambiguously, Johnston and Gilman (1959) compared microscopic plastic resistances with macroscopic plastic flow curves. This was done to see whether the micro and macro stresses are equal to each other. The next two figures show the results.



(a) (b)

Fig. 3.—Dislocation velocity *v*. stress for crystals that were given various treatments. (*a*) Effect of decreasing the temperature ; (*b*) effect of radiation damage.

Figure 3 shows dislocation velocity-stress curves for crystals that were given various treatments in order to change their hardnesses. It may be seen that the main effect of changing the " hardness " of a crystal is simply to shift the position of the dislocation velocity-stress curve along the stress coordinate. The amounts of shift do in fact equal the respective changes in the yield stresses, as Figure 4 shows. In this latter figure the stress that is needed to cause dislocation motion at a rate of $\sim 1\ \mu$/sec is plotted against a measure of the macroscopic flow stress.* It may be seen that the two quantities are proportional to each other. Therefore, it is concluded that the flow stress is determined by the plastic resistance of the crystal.

---

* The critical bending parameter is obtained from bending moment-deflection curves made in bend tests. It is the value of the bending moment at which plastic flow begins, divided by the section modulus.

If, in general, some other factor than the plastic resistance determined the flow stress, it would be difficult to understand why crystal hardnesses vary in a systematic way with lattice structure and lattice energy. MgO is always much harder than KI, for example, no matter how the crystals are prepared. If some critical length were important, such as the length of a Frank-Read source or the distance between the trees of a dislocation forest, then the flow stress ought to depend more sensitively than it does on the way in which a crystal is prepared.

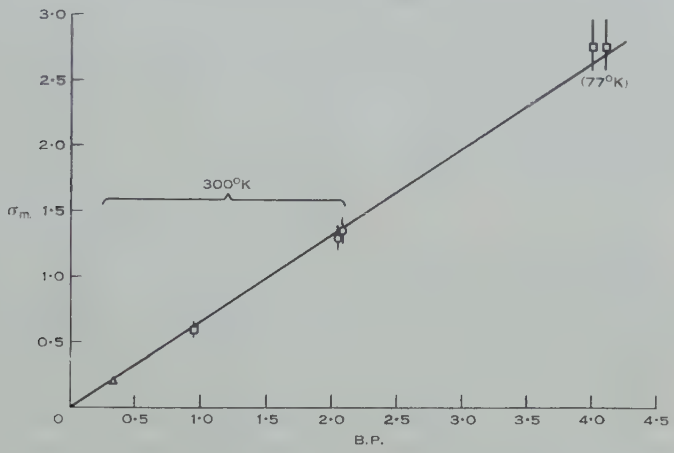


Fig. 4.—Stress needed to cause dislocation motion at a rate of ~1 μ/sec plotted against a measure of the macroscopic flow stress. *B.P.* = bending parameter = $M^*(c/I)$, σ_m = stress to move to a new dislocation. □, As grown (slowly cooled); △, cooled from 500 °C in 10 min; ○, neutron irradiated.

Since there would be a strong correlation of yield stress with the dislocation structures of crystals, there would be little reason to expect a correlation with lattice structure. This has been pointed out previously by Orowan (1954). Furthermore, in the particular case of LiF crystals, the initial dislocation density ($5 \times 10^4/\text{cm}^2$) is far too small to be able to determine the flow stress via the forest mechanism or through the internal stress fields of the initial dislocations.

Another factor which could be important is the impurity contents of crystals. The impurities might control the flow stress by forming Cottrell atmospheres around the dislocations in the crystals. However, this mechanism can only be important in special cases. Perhaps the simplest thing that it contradicts is the fact that the flow stresses of crystals often extrapolate to a finite strength for zero impurity content. Also, this mechanism is essentially a static one and so could be effective only at the beginning of flow. Actually, the flow stress is often nearly the same for a rapidly flowing crystal as for one that is just beginning to flow.

The dynamic nature of the hardness of crystals has been demonstrated directly for LiF crystals (Johnston and Gilman 1959). This was done by measuring dislocation velocities in crystals that had been hardened various amounts through neutron bombardments (Gilman and Johnston 1958). It was found that increased stresses are required to move dislocations in hardened crystals at the same velocities as in softer unirradiated ones. Most significant for this discussion is the fact that the increments of additional stress are almost independent of dislocation velocity. That is, the hardening is almost the same for dislocations moving at high speed ($\sim 5 \times 10^4$ cm/sec) as for dislocations moving slowly ($\sim 10^{-5}$ cm/sec). Clearly then, static pinning cannot account for the hardening. This behaviour is not restricted to irradiated crystals but appears to be the way in which hard versus soft crystals behave in general.

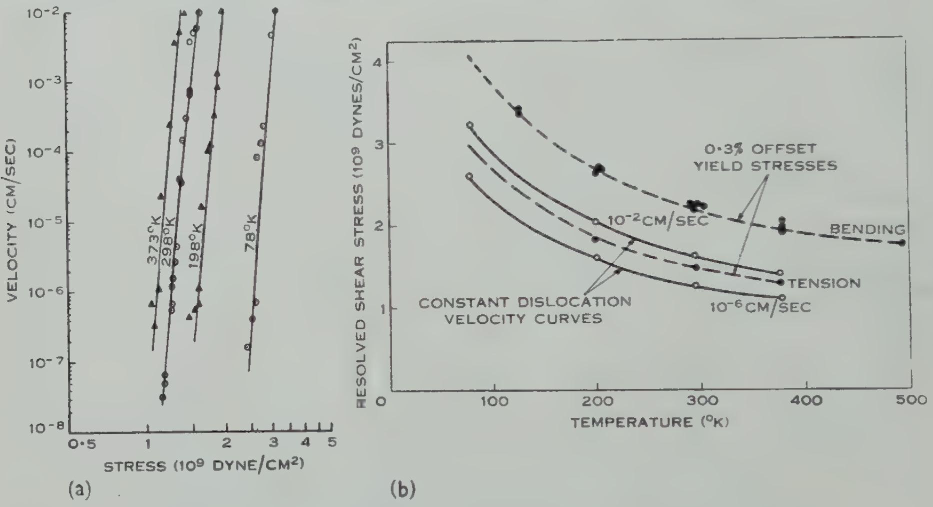


Fig. 5.—Edge dislocation motions in Fe-3% Si crystals (Stein and Low).

Lithium fluoride crystals do not behave in a manner that is unique to them. Recently, Stein and Low (1960) have made measurements of dislocation velocities in iron-silicon crystals (3 per cent. Si in Fe). They found that these crystals exhibit plastic resistance and behave qualitatively much like LiF crystals. Figure 5 shows some of their results. In Figure 5 (*b*) data are plotted which show that there is a definite correlation between the yield stresses of iron-silicon crystals and the stress that is needed to make dislocations move in them at a certain velocity. Less direct evidence of viscous dislocation motion in iron has been given by Cracknell and Petch (1955). Thus, in two rather different types of crystals, there is now direct evidence that plastic resistance controls plastic flow. Some evidence that this is true for all crystals will be presented in the next section of this paper.

III. QUESTION TWO: WHAT PROPERTIES OF A CRYSTAL DETERMINE ITS PLASTIC RESISTANCE?

The most important mechanical characteristic of a crystal is its elastic modulus. The role of the modulus in determining plastic resistance has sometimes been obscured because of complicating factors such as impurities, microstructure, and temperature, but data will be presented here which suggest that it plays a prime role.

The second most important feature of a crystal appears to be its electronic structure or chemical bond type. This is intimately connected, of course, with the crystal structure. In order to separate its effect from the effect of the elastic modulus one must therefore consider each structural group of crystals separately.

One extreme structural group is the face-centred cubic metals. Dislocation velocity data are not available for these crystals, so, in order to compare them, it was necessary to use hardness data as a measure of plastic resistance. This is a reasonable way of making comparisons since modern microhardness machines can make permanent indentations in almost all crystals without causing cracking ; therefore they measure a number that is a measure of the yield stress of a crystal. In fact, for many materials, the hardness number has been shown to be directly proportional to the yield stress. Since the yield stress is proportional to the plastic resistance (Fig. 4), this means that the indentation hardness is also proportional to the plastic resistance.

Data for the Young's moduli of face-centred cubic metals at −200 °C were collected from the work of Koster (1949). Brinell hardness data came from other sources (Winkler 1943 ; Druyvesteyn 1947 ; American Society for Metals 1948). The conversion from the Young's modulus E to the shear modulus G is simple, since $G=3/8E$ is a good approximation (Zwikker 1954). The hardness numbers can be converted to yield stresses with confidence, since the correlation between these quantities is quite well understood (Tabor 1948). The yield stresses that are obtained in this way are representative of polycrystalline samples, but they can be expected to be proportional to the yield stresses of mono-crystals.

The face-centred cubic metals are particularly suited to the type of correlation that is attempted here for two reasons. First, being noble metals, they are readily purified so that measurements on quite pure samples are available. Second, the dependence of their yield stresses on impurity content, and on temperature, is quite small and hence large errors due to these factors are unlikely.

The excellent correlation that exists between yield stress and elastic modulus for face-centred cubic metals is shown in Figure 6. The simple proportionality of the two quantities strongly indicates that plastic resistance determines the stress needed for plastic flow in these metals as well as in LiF and Fe-Si. If the flow stress depended on geometric properties of the dislocation networks of the crystals, the relative dimensions of such networks would need to be constant for the whole series of crystals in order to get the correlation of Figure 6. This seems highly unlikely.

A similar correlation for body-centred cubic metals may be seen in Figure 7. Here the data are considerably less reliable than in the case of face-centred cubic metals. None the less there does seem to be a relation between the elastic moduli and plastic flow stresses of these metals. With the exception of the K and Na points, the flow stress data are from sources given by Bechtold (1959). They have been measured on the purest available metals. Since the yield stresses of body-centred cubic metals are very sensitive to impurities and temperature, the

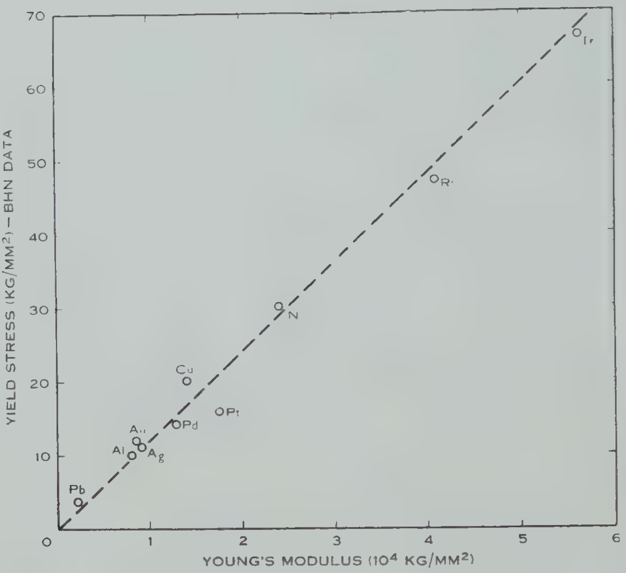


Fig. 6.—Relation of yield stress to elastic modulus for face-centred cubic metals (at —200 °C). $\sigma_y = 1 \cdot 2 \times 10^{-3} E$, $\tau_y = 1 \cdot 6 \times 10^{-3} G$.

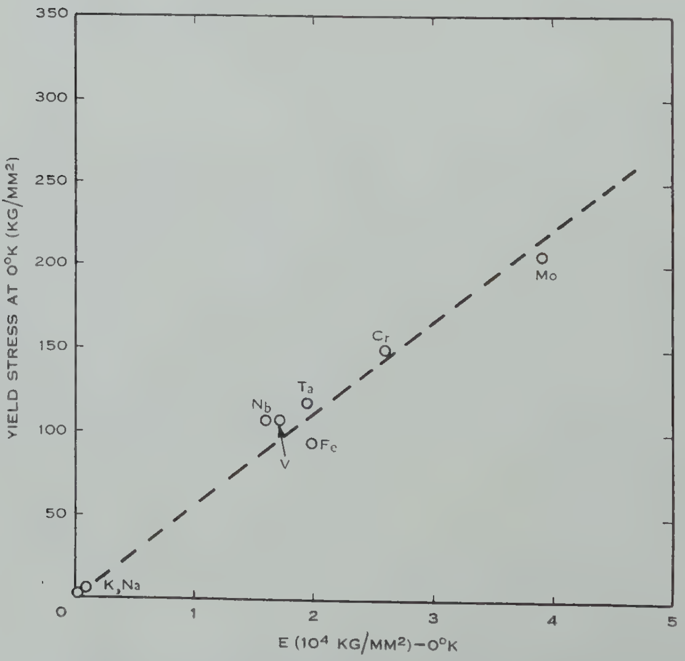


Fig. 7.—Relation of low temperature yield stress to elastic modulus for body-centred cubic metals. $\sigma_y = 5 \cdot 6 \times 10^{-3} E$, $\tau_y = 7 \cdot 5 \times 10^{-3} G$.

stresses were extrapolated to 0 °K on a log yield-stress versus temperature plot, and the 0 °K values are shown in Figure 7. Bechtold's stresses for tungsten seemed to be too high because of impurities, so they were omitted. The K and Na points are estimates from room-temperature hardness data (Edwards 1918).

Covalent crystals behave in extreme contrast to metal crystals. They are typified by crystals with the zinc blende structure, and data for them are shown in Figure 8. The hardness data come from various sources including Tabor (1954), Westbrook (unpublished data), and Borschevsky (1958); and the elastic constants were taken from Huntington's (1958) review. The hardness values for diamond appear to be somewhat high; the values are in some doubt because the indenters that are used for the measurements are made of diamond itself. If a harder indenter were used it is likely that diamond would appear to be somewhat softer. The fact that some of the soft compounds do not fall on the trend line in Figure 8 may result because the temperature of measurement (300 °K) is an appreciable fraction of their melting points.

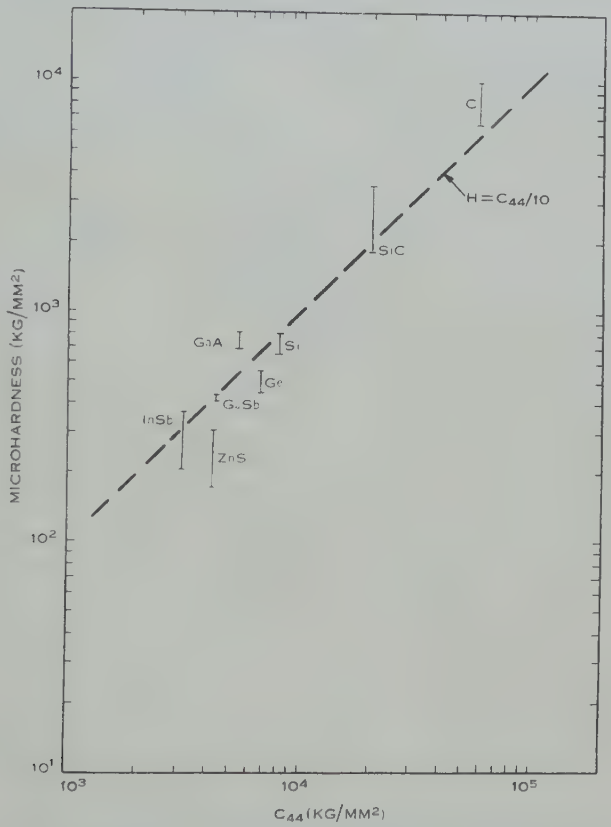


Fig. 8.—Hardness *v.* elastic modulus for crystals with the zinc blende structure.

The main feature of the data for the covalent compounds is the high ratio of H/C_{44} in comparison with the H/G ratio for metals. The fact that the hardnesses of these compounds correlate quite well with their elastic shear constants suggests once more that plastic resistance controls their yield stresses. However, their high H/C_{44} ratios indicate that dislocations have great difficulty in moving through the crystals at low temperatures. Since H/C_{44} is not greatly sensitive to impurities, the reluctance of dislocations to move in covalent crystals is considered to be one of their intrinsic properties.

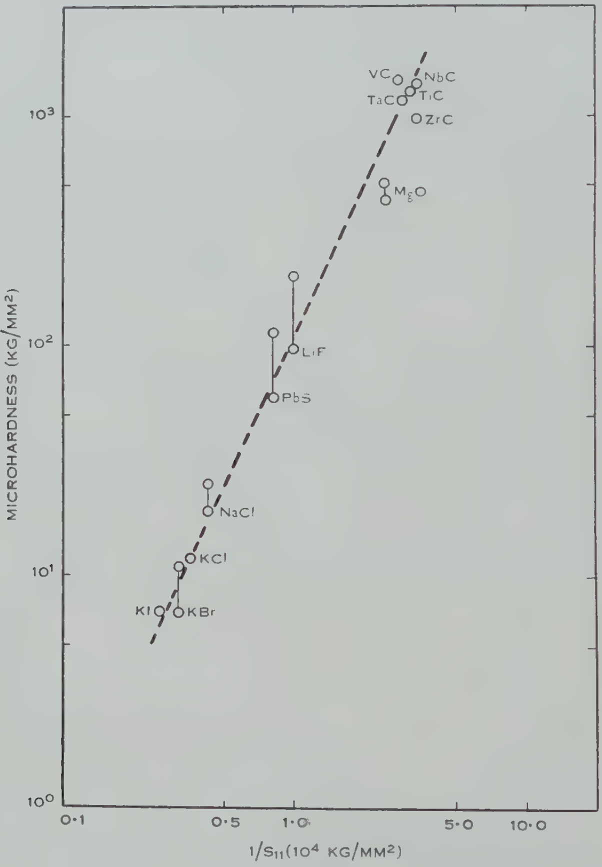


Fig. 9.—Relation between microhardness and elastic modulus for crystals with the rock-salt structure.

Ionic crystals behave in a manner that is different than that for metallic and covalent ones. This may be seen in Figure 9, which is plotted from data of Westbrook (unpublished data) and Bennhardt (1941). In this case the hardness is not simply proportional to the elastic modulus but appears to depend on some power higher than unity. Perhaps the point is even more clear in Figure 10 which shows the trend lines for all three types of crystal. It appears that the softer crystals with the rock-salt structure resemble metals in their behaviour, while the hardest ones approach the behaviour of covalent crystals. The trend line would shift somewhat, however, if all the data were for low homologous temperatures. Some of the points in Figure 9 were measured at relatively high homologous temperatures.

It has been known for some time, of course, that covalent crystals tend to be harder than metallic crystals. The purpose of the plots presented here is to show that the difference is roughly constant if the effect of the elastic modulus is factored out. Also, since the hardness correlates so well with the elastic modulus it appears to be related to an intrinsic property of each crystal, namely, the plastic resistance of the crystal.

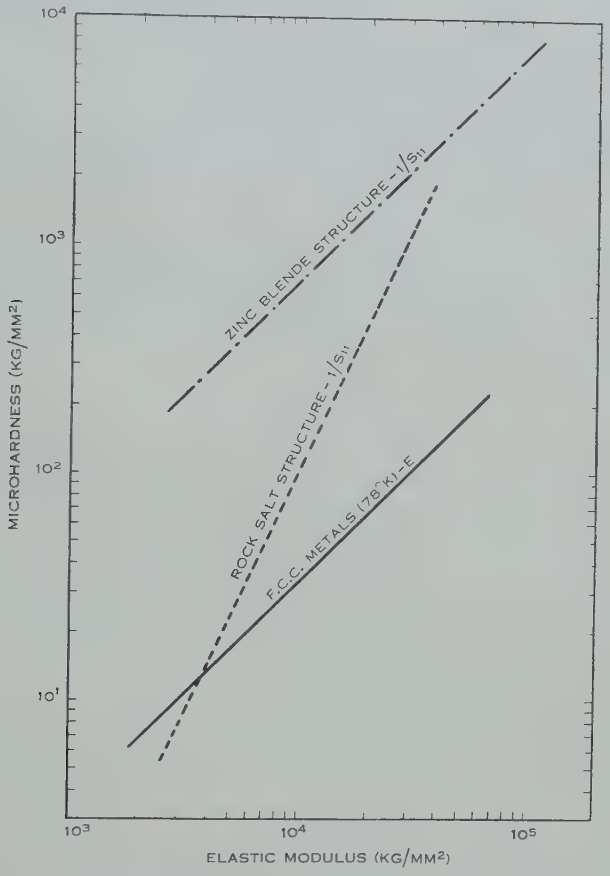


Fig. 10.—Hardness *v.* elastic modulus for crystals with the same lattice structures.

The differences in behaviour of different structural types (Fig. 10) do not appear to be closely related to differences of the Burgers vectors of glide dislocations. Within a given structural group, crystals with the same Burgers vectors may have quite different hardnesses. For example, MgO and LiF have Burgers vectors of 2·97 and 2·85 Å respectively, yet MgO is substantially harder than LiF. Also, Si and ZnS both have Burgers vectors of 3·8 Å, but quite different hardnesses. Similarly, one may compare different structural types with the same Burgers vectors. NaCl and Ge both have Burgers vectors of 4·0 Å, but Ge is more than 50 times as hard as NaCl; another case is MgO and SiC, which have Burgers vectors of 3·0 and 3·1 Å, respectively.

The last example above, that of MgO and SiC, is especially interesting because not only the Burgers vectors but also the elastic moduli of these two crystals are quite similar; C_{44} is equal to $1 \cdot 5 \times 10^3$ kg/mm^2 for MgO and $\sim 2 \cdot 0 \times 10^3$ kg/mm^2 for SiC. A more extreme example is that of Ir and SiC. For these crystals $b = 2 \cdot 7$ and $3 \cdot 1$ respectively while $G = 2 \cdot 1$ and $2 \cdot 0 \times 10^3$ kg/mm^2, respectively; yet SiC is 35 times as hard as Ir. Still another comparison might be Ni and VC for which $b = 2 \cdot 5$ and $3 \cdot 0$ Å, respectively; and $E = 2 \cdot 4$ and $2 \cdot 8 \times 10^3$ kg/mm^2, respectively. The hardness values are 30 for Ni and 1450 kg/mm^2 for VC; a factor of 50 different.

It may be seen from the above discussion that the elastic energies of dislocations do not play the dominant role in determining the yield stress. The elastic energy is proportional to Gb^2, but quite different hardnesses are observed even though Gb^2 is kept constant.

Since the hardness is a measure of the rate at which moving dislocations dissipate energy, it must be concluded that dislocations dissipate considerably more energy as they move through covalent crystals than when they move through metal crystals. The reason for the greater rate of energy dissipation is not known at present.

(*a*) *Dislocation Structure*

The importance of the details of atomic binding in determining dislocation mobilities may be demonstrated by the influence of the glide plane on dislocation mobility. It is possible to apply stresses to some crystals in such a way that the shear stress on the primary glide plane is zero. Then the crystal is forced to glide on some other plane. The glide usually occurs in the primary glide direction so the only factor that is changed is the glide plane. If specimens from the same crystal are used, the effects of factors such as the impurity content and the initial dislocation structure can be eliminated.

One case where the glide plane has a large effect is in zinc crystals. This case has been studied by Gilman (1956). He pulled on zinc crystals parallel to their basal planes in order to force them to glide on prism planes. Very large stresses are needed to make zinc crystals flow at a given strain rate on their prismatic planes $\{10\bar{1}0\}$ as compared with their basal planes (0001). This is shown in Figure 11. The glide direction $\langle 12\bar{1}0 \rangle$ was the same for the two types of flow. Beryllium shows a similar anisotropy in plastic resistance (Tuer and Kaufman 1955).

Crystals with the rock-salt structure usually glide on {110} planes, but they can be forced to glide on {100} planes if the temperature is high enough (Gilman 1959). Data that show this effect for LiF crystals are plotted in Figure 12. Similar behaviour has been observed in NaCl crystals (Dommerich 1934) as well as KCl, KBr, and KI (Gilman 1959). This case provides convincing evidence of the importance of the structure of dislocation cores on mobility because the Burgers vector is the same for both planes, and unlike the case of zinc the elastic constants are nearly the same for the two planes. Hence, the elastic energies of the dislocations are nearly equal, leaving only the core structures to cause the difference in behaviour.

Anisotropies in the plastic resistances of crystals are probably related to the behaviour of edge dislocations since screw dislocations have less distinct glide planes. Some evidence of this is found in LiF crystals where screw dislocations sometimes move on {100} planes at room temperature (Gilman and Johnston 1957), but full-scale {100} glide is only observed at temperatures above about 175 °C (Fig. 12).

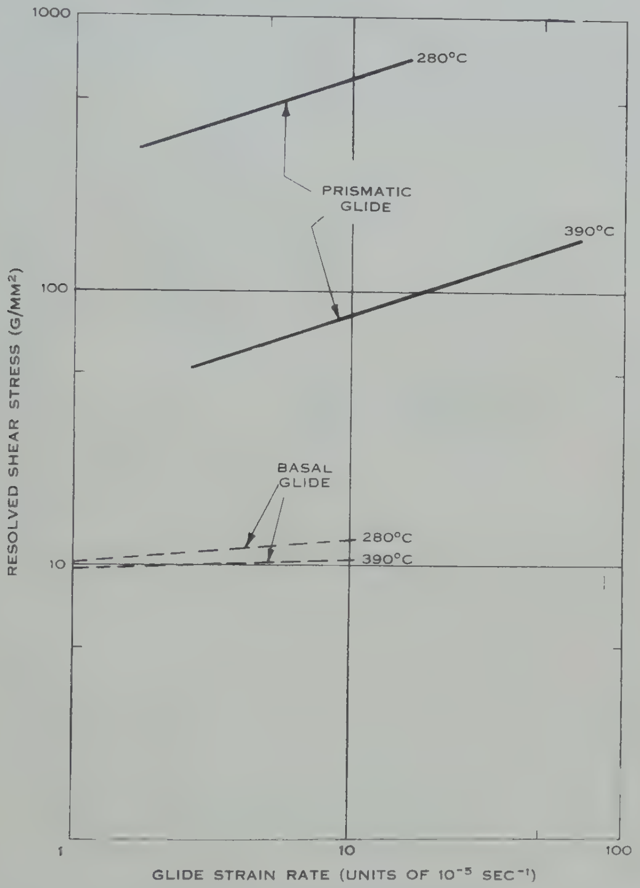


Fig. 11.—Comparison of stresses required to cause basal versus prismatic glide in zinc crystals.

(b) *Internal Stresses*

The fact that the main effect of changing the " hardness " of a crystal is to shift its dislocation velocity-stress curve on the stress coordinate suggests that internal stresses are the cause of the change of hardness (Gilman and Johnston 1958). The effects are additive, so it seems that pure crystals of a given structural type have an intrinsic velocity-stress curve which is a function of temperature. Fluctuating internal stresses of a certain average amplitude slow down and speed up moving dislocations which encounter them. This causes a net decrease in the average rate of motion. Additional applied stress is then needed to make them move at the velocity characteristic of a pure and relatively perfect crystal. This additional stress is just equal to the average internal stress amplitude (Gilman and Johnston 1958).

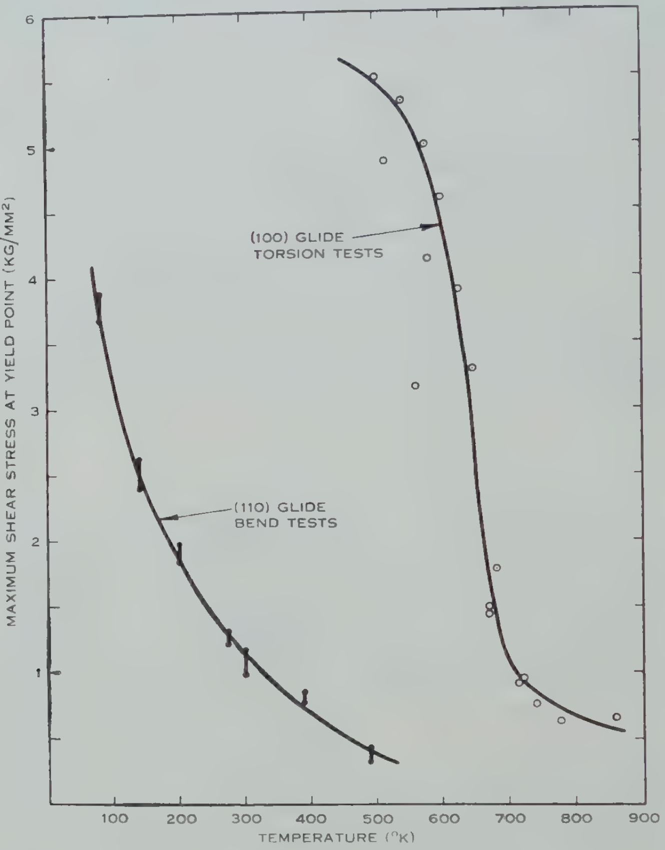


Fig. 12.—Comparison of stress for (100) glide with stress for (110) glide in LiF crystals.

IV. Question Three: How does Temperature Affect the Plastic Resistances of Crystals?

The mobility of dislocations in face-centred cubic metal crystals decreases very slowly with increasing temperature, but in other crystals very large changes occur; especially in covalent crystals. Some examples are shown in Figure 13 (Patel and Alexander 1956; Wachtman and Maxwell 1957; Lillie 1958).

There are two ways of considering the variation of plastic resistance with temperature. One possibility is that dislocations are pinned somehow at low temperatures (say by the Peierls-Nabarro force) and thermal waves aid the applied stress in breaking them loose from their pinned positions. This possibility seems unlikely because the variation of average dislocation velocity with applied stress would be linear for small stresses. Something quite different is observed (Fig. 2). A second possibility is that the pinning force itself decreases with increasing temperature. Thus the variation with stress for stresses above the yield stress would not depend strongly on temperature but the value of the yield stress would. This is what is observed.

It is suggested that a dislocation in a covalent crystal has a " melting point ". By this it is meant that the plastic resistance or viscosity at the core of a dislocation suddenly changes from a high value to a much lower value as the temperature increases. This is analogous to the general loss of viscosity in a glass or crystal at its melting point. Perhaps the analogy with the loss of viscosity in a glass is best. Thus the dislocation core may be considered to be like a rigid supercooled liquid at low temperatures which loses its rigidity at elevated temperatures. Dislocation cores in face-centred cubic metals might also be considered to be like a small tube of liquid, but for them the temperature never becomes low enough to make the liquid rigid. The criterion that determines when the " melting temperature " is reached may be a critical r.m.s. atomic vibration amplitude or a critical amount of " free volume " at the dislocation.

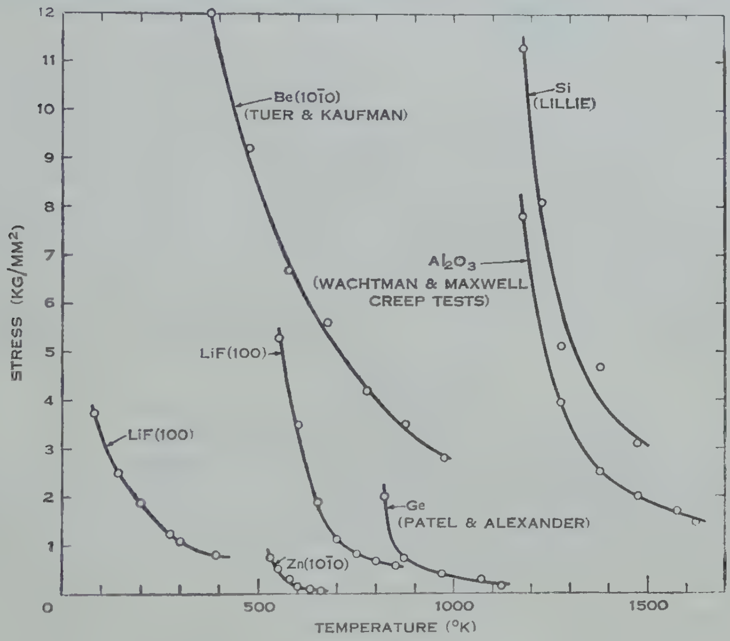


Fig. 13.—Effect of temperature on the flow stresses of crystals that have some covalent binding.

When detailed data on the behaviour of dislocations in covalent crystals become available, it may be worth while to attempt to treat this model quantitatively, but this does not seem appropriate at the present time.

V. Conclusions

It has been pointed out that the most important factor that determines the yield or flow stress of a crystal is the *plastic resistance* of the crystal to dislocation motion. This has been demonstrated directly in the case of LiF crystals by

Johnston and Gilman, and more recently in Fe-Si crystals by Stein and Low. In both the ionic and the metallic crystals, the plastic resistance is almost independent of dislocation velocity except at very high velocities. Therefore, the damping force is almost independent of dislocation velocity for low velocities, so the viscosity is non-Newtonian.

The core structures of dislocations are important in determining the plastic resistance to their motion. This is proved by the fact that the glide plane of a moving dislocation sometimes has a large effect on the plastic resistance. Examples are zinc and alkali halide crystals. For the same glide direction, zinc glides much more readily on the basal planes of the hexagonal structure than on the prismatic planes. Also, LiF and NaCl glide more easily on {110} planes than on {100} planes.

Internal stresses in crystals caused by impurities, radiation damage, etc. simply add to or subtract from the applied stress. In this way they cause a net reduction of the average dislocation velocity for a given applied stress.

Hardness data for many crystals are presented as a function of their elastic moduli. Provided that crystals with the same lattice structure are considered, the hardness data correlate very well with the elastic moduli. For pure face-centred cubic metals at −200 °C the yield stress is proportional to the shear modulus ; $\tau_y = 1 \cdot 6 \times 10^{-3} G$. For covalent crystals with the zinc blende structure the hardness is approximately equal to $C_{44}/10$. Ionic crystals with the rock-salt structure behave in an intermediate fashion.

In view of these data, it is believed that for all crystals the plastic resistance to dislocation motion is what determines the yield stress.

The plastic resistances of covalent crystals are very large at low temperatures and then decrease suddenly at temperatures equal to about two-thirds of the melting point. It is suggested that this is because the cores of dislocations in covalent crystals " melt " at this temperature.

VI. ACKNOWLEDGMENTS

I am indebted to several of my associates at the General Electric Research Laboratory. Dr. J. H. Westbrook provided me with some of his unpublished hardness results for rock salt type crystals and suggested many sources of data. Dr. J. R. Low allowed me to use data obtained by Mr. D. L. Stein and himself prior to its publication. Dr. D. Turnbull provided helpful discussion.

VII. REFERENCES

AMERICAN SOCIETY FOR METALS (1948).—" Metals Handbook." (A.S.M. : Cleveland, Ohio.)

BECHTOLD, J. H. (1959).—Cleavage in the refractory metals. Swampscott Fracture Conference 1959.

BENNHARDT, E. O. (1941).—*Z. Metallk.* **33** : 135.

BORSCHEVSKY, A. S., GORUNOVA, N. A., and TAKMAREVA, N. K. (1958).—*J. Tech. Phys., Moscow* **27** : 214.

COTTRELL, A. H. (1953).—" Dislocations and Plastic Flow in Crystals." p. 17. (Oxford Univ. Press.)

CRACKNELL, A., and PETCH, N. J. (1955).—*Acta Metall.* **3** : 186.

DOMMERICH, S. (1934).—*Z. Phys.* **90** : 189.

DRUYVESTEYN, M. J. (1947).—*Appl. Sci. Res., Hague* A **1** : 66.

Edwards, C. A. (1918).—*J. Inst. Met.* **20** : 86.

Fisher, J. C. (1955).—*Trans. Amer. Soc. Metals* **47** : 451.

Gilman, J. J. (1956).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **206** : 1326.

Gilman, J. J. (1959).—Plastic anisotropy of LiF and other rock salt type crystals. *Acta Metall.* **7** : 608.

Gilman, J. J., and Johnston, W. G. (1957).—"Dislocations and Mechanical Properties of Crystals." p. 116. (Wiley : New York.)

Gilman, J. J., and Johnston, W. G. (1958).—*J. Appl. Phys.* **29** : 877.

Huntington, H. B. (1958).—*Solid State Phys.* **7** : 214.

Johnston, W. G., and Gilman, J. J. (1959).—*J. Appl. Phys.* **30** : 129.

Koster, W. (1949).—*Z. Metallk.* **39** : 1.

Lillie, D. W. (1958).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **212** : 249.

Orowan, E. (1954).—"Dislocations in Metals." p. 125. (A.I.M.E. : New York.)

Patel, J. R., and Alexander, B. H. (1956).—*Acta Metall.* **4** : 385.

Stein, D. L., and Low, J. R. (1960).—*J. Appl. Phys.* **31** : 362.

Tabor, D. (1948).—*Proc. Roy. Soc.* A **192** : 247.

Tabor, D. (1954).—*Proc. Phys. Soc.* B **67** : 249.

Tuer, G. L., and Kaufman, A. R. (1955).—"The Metal Beryllium." p. 372. (A.S.M. : Cleveland, Ohio.)

Wachtman, J. B., and Maxwell, L. H. (1957).—*J. Amer. Ceram. Soc.* **40** : 377.

Winkler, O. (1943).—*Z. Elektrochem.* **49** : 221.

Zwikker, C. (1954).—"Physical Properties of Solid Materials." p. 90. (Interscience : London.)

Appendix I

Theory of Viscous Dislocation Motion

Perhaps the most unusual feature of the data of Figure 2 is the great sensitivity of the dislocation velocities to the applied stress. This rather severely restricts the possible types of theory that can be used to describe the data. No complete theory is available at present, but it may be worth while to present a simple theory that is semi-quantitatively in agreement with the observations. The form of the experimental data suggests that the stress dependence must be of the exponential type ; $\bar{v}=v_0 \exp(A/\tau)$, where v is the dislocation velocity, τ is the shear stress, A and v_0 are constants. Then a plot of log $\bar{v}$ versus reciprocal τ should yield a straight line. Such a plot is shown in Figure 14 for the data on screw dislocations, and indeed the data follow a straight line whose equation is

$$\bar{v}=10^5 \exp(-16\cdot 8\times 10^8/\tau),$$

where the velocity has the dimensions of cm/sec and the stress is given in dynes/cm².

The theoretical interpretation of this result is based on a paper by J. C. Fisher (1955). We suppose that a dislocation line lies initially in a position in a crystal such that its potential energy per unit length, P_0, is less than the value P_R that it would have for a random position in the crystal. The reduction of potential energy could come from a variety of causes. It might be due to the Peierls-Nabarro energy or to the Cottrell interaction energy, but, whatever the cause, it is localized to atomic dimensions. If a small piece of the dislocation line moves away from the initial position the situation looks schematically like

Figure 15. The motion of the segment BC of the dislocation line changes the energy of the system in three ways:

(*a*) the length of dislocation line is increased;
(*b*) the energy of the segment per unit length is greater in its final position than in its initial position;
(*c*) some plastic work is done when the dislocation glides over the cross-hatched area.

The energy of the segment in its free position is $2r\theta P_R$, whereas its energy in the bound position is $2rP_0 \sin\theta$, and the work that is done by its gliding is the force on the dislocation, τb, times the area that is swept out, $\theta r^2 - r^2 \sin\theta \cos\theta$.

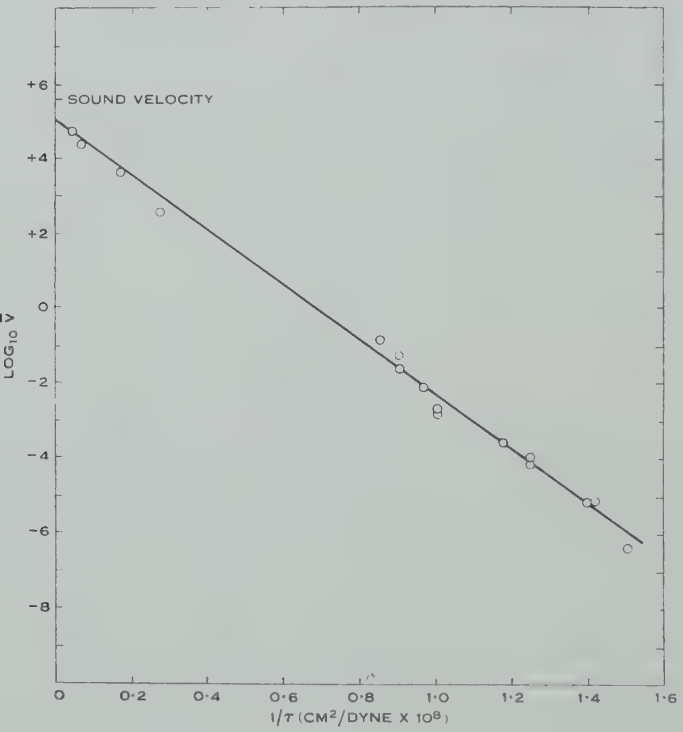


Fig. 14.—Screw dislocation velocities in a LiF crystal. The equation of the straight line is $v = 10^5 \exp(-16\cdot8\times10^8/\tau)$, where v is in cm/sec and τ is in dynes/cm^2.

It is assumed that the arc is free to equilibrate with the applied stress at all times so that $r = P_R/b\tau$. Then the work of forming the loop will be

$$W = [2\theta P_R - 2P_0 \sin\theta - P_R(\theta - \sin\theta\cos\theta)]P_R/b\tau.$$

The critical value of W for successful nucleation of the motion is obtained when $\partial W/\partial\theta = 0$; that is, when $\cos\theta^* = P_0/P_R$. If this is substituted into the equation for W, the critical work of formation is obtained

$$W^* = \frac{P_R^2}{b\tau} f(P_0/P_R),$$

where $f(P_0/P_R) \simeq 2/3[2(1-P_0/P_R)]^{3/2}$ for $P_0/P_R \approx 1$. The probability that such a nucleus will form is proportional to $e^{-W^*/kT}$, and each time a nucleus forms it allows the dislocation to advance a distance of the order of b. Hence, if we call the atomic frequency factor ν, the dislocation velocity is given approximately by

$$v = b\nu e^{-W^*/kT}.$$

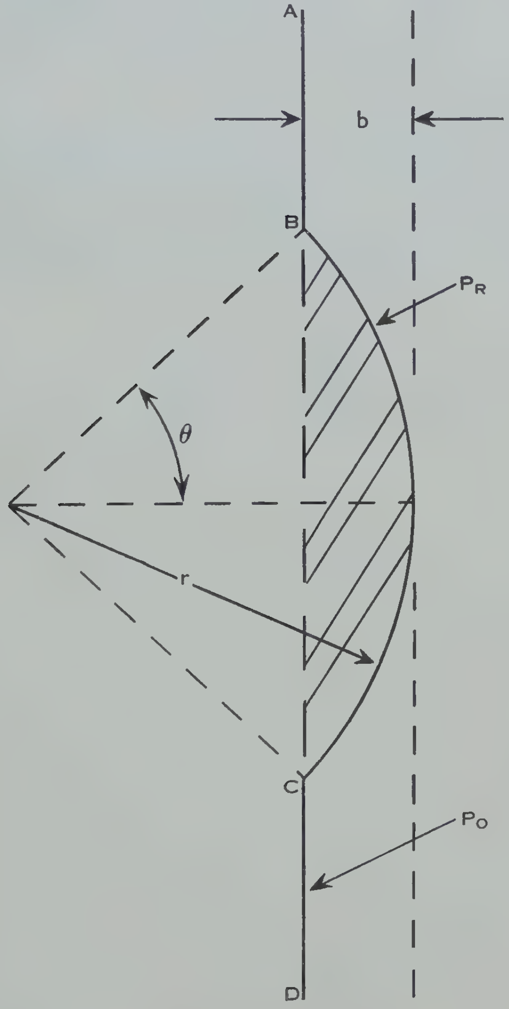


Fig. 15.—Nucleation of a dislocation movement. The initial position is the line *AD* and the final position is the heavy solid line.

Comparing this with the experimental equation given above we see that

$$W^*/kT = 16 \cdot 8 \times 10^8/\tau.$$

Then, since P_R for LiF is about 2×10^{-4} ergs/cm, and $b = 2 \cdot 85$ Å; at $T = 300$ °K we find that $f(P_0/P_R) \simeq 5 \times 10^{-5}$; so $P_0/P_R \simeq 0 \cdot 999$. This is in accord with our assumption, and it is to be expected because most of a dislocation's energy is independent of the energy at its core.

The experimental value of $b\nu$ is 10^5 cm/sec, so $\nu = 3 \cdot 5 \times 10^{12}$ sec; a value that seems reasonable.

Another numerical comparison can be made with Johnston and Gilman's measurement of the temperature dependence of dislocation motion in LiF. At a stress of $1 \cdot 1 \times 10^8$ dynes/cm^2 they measured an effective activation energy of about $0 \cdot 7$ eV. The equation given above yields an activation energy for the same conditions of about $0 \cdot 4$ eV. This is fairly good agreement considering the crude assumptions of the theory.

If a theory of this form is correct, the temperature dependence of the flow will be quite complicated because it is not expected that P_0/P_R will be independent of temperature. In fact, if the viscosities of dislocations decrease abruptly with increasing temperature, as was suggested in the paper, then P_0/P_R will be less than unity at low temperatures but nearly equal to unity at high temperatures.

ANNEALING KINETICS OF NEUTRON-IRRADIATED ALUMINIUM AND COPPER

By T. H. Blewitt,* R. R. Coltman,* and C. E. Klabunde*

[*Manuscript received October 8, 1959*]

Summary

Activation energies for the annealing of copper and aluminium following reactor bombardment near 4 °K have been measured in the range from 10 to 40 °K. Both the change in slope method and the isothermal technique method were utilized with the assumption that a constant activation energy existed. Computations of the number of jumps involved from the measured activation energy result in an impossibly small number. It is obvious that the method for determination of the activation energies is not applicable, probably because of the non-uniqueness of the activation energy.

I. Introduction

The earliest concepts of radiation damage in metals assumed that fast neutrons interacted with atoms in their equilibrium position, resulting in their displacement and the formation of vacant lattice sites and interstitial atoms. Since theoretical considerations indicated that interstitials might have a significant mobility at low temperatures (Huntington 1953), considerable effort has been spent both here and at other laboratories in the study of metals bombarded at low temperatures. As a function of time the bombardment temperatures have been getting progressively lower, and at present the lowest reported bombardment temperature is 3·8 °K. These experiments have shown that some defect is indeed mobile at low temperatures, and in the case of aluminium and copper a significant annealing peak has been observed in the temperature interval from 30 to 50 °K. The earliest interpretations assumed that this annealing peak was associated with a migration of interstitial atoms to vacant lattice sites. One of the most difficult problems arising in this interpretation lies in the fact that only about half of the defects appear to be annihilated. Since equal numbers of vacancies and interstitials are created, it is difficult to imagine why only half of them would disappear if the above hypothesis were valid. The experiment discussed here was devoted to a study of the annealing kinetics in the region below 50 °K. In particular, the experiments are designed to test the validity of interstitial migration and annihilation in this temperature interval.

II. Experimental Techniques

Bombardments were made in the Hole 12 Cryostat of the Oak Ridge Graphite Reactor. This facility, which incorporated a liquefier attachment, has been previously described (Coltman, Blewitt, and Noggle 1957 ; Howe, Coltman, and Blewitt 1957). Bombardment temperatures were in the region between 3·8

---

* Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

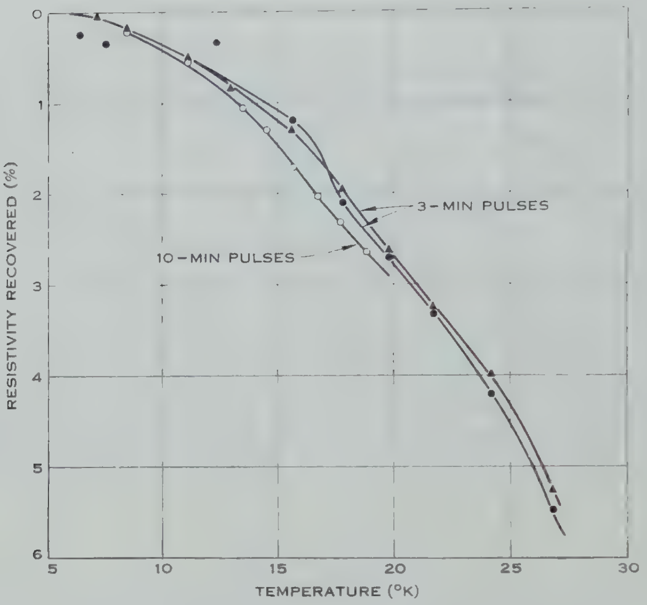


Fig. 1.—Isochronal pulse annealing of copper following 3·9 °K reactor bombardment.

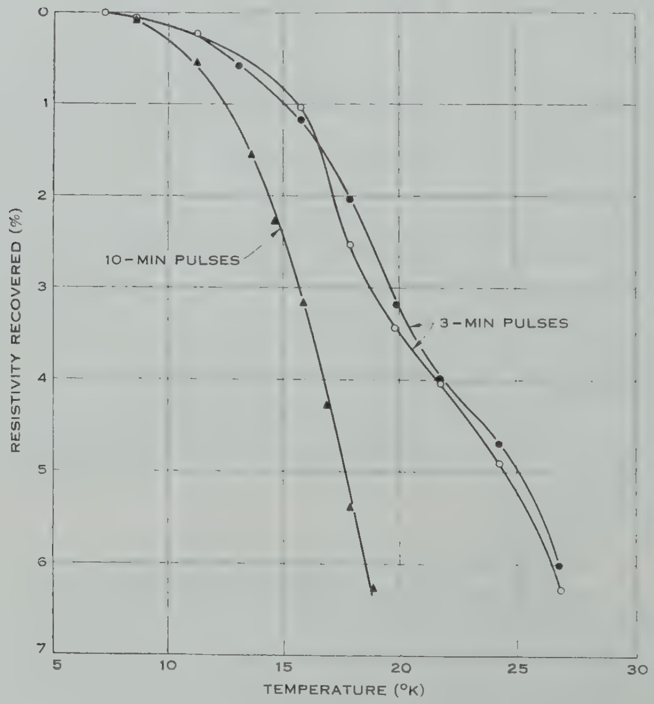


Fig. 2.—Isochronal pulse annealing of aluminium following 3·9 °K reactor bombardment.

and 4·2 °K. Annealing was performed by locating the samples in a vacuum-tight can, together with a resistance heater. To pulse the samples, the sample can was isolated from the wall of the cryostat and the heater utilized to heat the samples. The relatively low heat capacity at this temperature made it possible to raise the temperature of the samples as much as 30 °K in a matter of a few seconds. The samples were rapidly quenched by injecting helium exchange gas into the space between the sample can and the cryostat wall. The time required to quench the sample 10 or 15 °K was only a few seconds. By this technique both isothermal and isochronal anneals were made.

Isochronal anneals were made in the region from 4·2 to 40 °K. It might be noted that the temperature was measured by means of a copper-constantan thermocouple. It is estimated from the reproducibility of the isochronal annealing curves that the absolute temperature is not in error by more than 0·3 °K. Some difficulty is encountered in the measurements due to the presence of Thomson e.m.f.'s. It has been established, however, that the neutron irradiation does not affect the thermoelectric output of the couple by more than 1×10^{-8} V for the doses considered here. Furthermore, it is estimated that the relative temperatures are not in error by more than 0·1 °K.

III. RESULTS

The results of the isochronal annealing are shown in Figures 1 and 2. Here we have plotted the resistivity recovery as a function of pulse temperature. The measurements were all at the base temperature of 3·8 °K. It will be noted that in the first run (3-min pulses) for both aluminium and copper there was a dip at approximately 17 °K, somewhat suggestive of the type reported by Walker and co-workers (Corbett and Walker 1958). The fact that both metals show the same ambiguity makes it questionable that this is a real effect and, indeed, the subsequent measurements of Run 2 and Run 3 fail to show any such dip. Run 3 for both aluminium and copper (10-min pulses) is representative of the final points obtained during the study of isothermal annealing.

Isothermal annealing curves were obtained in approximately 2° intervals in the range from 9 to 20 °K. The primary purpose of these experiments was to determine the activation energy, utilizing the change in slope method suggested by Overhauser. At each temperature four points were determined. A typical plot of the isothermal annealing curves is shown in Figure 3. It is estimated that ratios of the slopes can be determined to within a factor of 2. Since this ratio appears as a logarithm, this uncertainty is not too serious. The computation of the activation energies from these data is shown in Table 1. The determination of activation energies in the region between 30 and 50 °K was also determined by the same method. The change in slope resulting when the temperature is raised from 36·2 to 38·9 °K is shown in Figure 4. The activation energy in this temperature range was also determined utilizing the relaxation time method. This is done by measuring the time required to anneal out a given fraction of the resistance at two different temperatures. The results obtained to determine the activation energy by this method are shown in Figure 5. The activation energies determined by both these techniques are shown in Table 2.

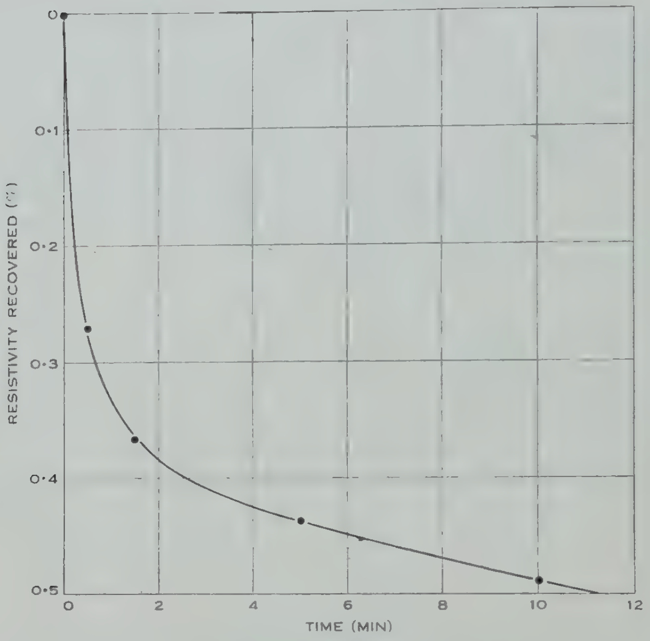


Fig. 3.—Isothermal pulse annealing of copper at 13·5 °K.

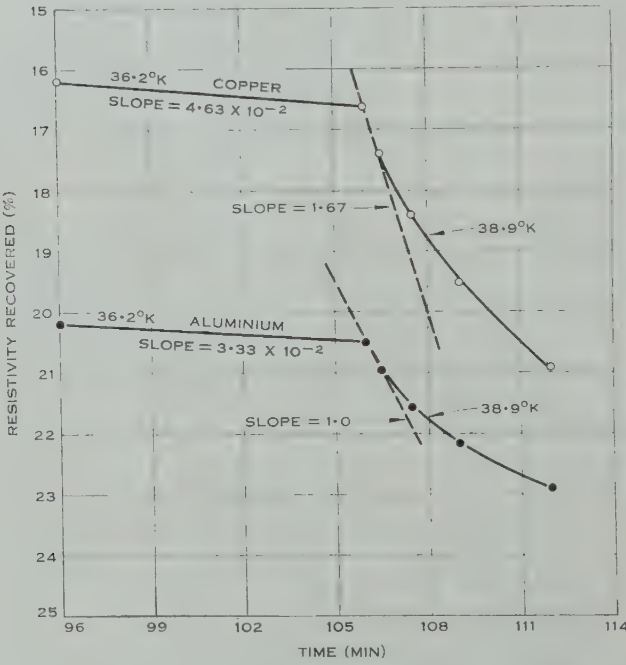


Fig. 4.—Isothermal pulse annealing of irradiated copper and aluminium.

TABLE 1

APPARENT ACTIVATION ENERGY FOR COPPER AND ALUMINIUM DETERMINED FROM THE CHANGE IN SLOPE METHOD

Temperature (°K)	Aluminium		Copper	
	ε (eV)	$\varepsilon_{\text{min.}}$* (eV)	ε (eV)	$\varepsilon_{\text{min.}}$* (eV)
8·5				
	0·020	0·020	0·022	0·019
11·2				
	0·032	0·032	0·035	0·031
13·5				
	0·054	0·049	0·040	0·039
14·5				
	0·049	0·049	0·065	0·065
15·7				
	0·081	0·069	0·080	0·068
16·75				
	0·077	0·070	0·095	0·088
17·8				

* $\varepsilon_{\text{min.}}$ is the minimum energy determined by drawing a straight line between the last point of isothermal annealing at T_i and the first point of the isothermal anneal at temperature T_{i+1}.

TABLE 2

APPARENT ACTIVATION ENERGY FOR RECOVERY OF ALUMINIUM AND COPPER AT 36·2 °K, 38·9 °K

Metal	Time at 38·9 °K (min)	Time at 36·2 °K (min)	Energy (eV)
Copper	0·47 0·97 1·64	20·5 59·7 104	0·171 0·186 0·187
	From change in slope		0·164
Aluminium	0·408 0·701 1·32 3·80	11·4 22·0 43·5 106·0	0·150 0·156 0·158 0·150
	From change in slope		0·153

IV. Discussion

The determination of the activation energies and of the isochronal annealing curves points out several interesting facts. First, it should be noted that annealing occurs in copper at temperatures as low as 7·2 °K. It would seem very difficult to ascribe this annealing to the same simple defect which is moved in the 30 to 50 °K range. Some difficulty would, therefore, be expected in explaining the whole low temperature annealing as the result of the motion of a simple defect. Secondly, it should be noted that the activation energies increase monotonically as a function of temperature. The discrete steps found by other investigators do not appear to be present here (Corbett and Walker 1958). It would thus seem difficult to apply the theory that close vacancy interstitial pairs of discrete activation energies were present to account for low temperature annealing. Finally, the magnitude of the activation energies seems much too high to allow a simple process to be occurring. This is particularly noted when one calculates the number of jumps required for annihilation for any of the activation energies which have been determined. The results of such computations are shown in Table 3. For example, the activation energy of 0·187 eV found at 37 °K for copper would require 10^{-10} jumps. It is readily seen that this result is absurd, and one is forced to conclude that the rate equation cannot be applied to these results. It would seem that the explanation for this anomaly

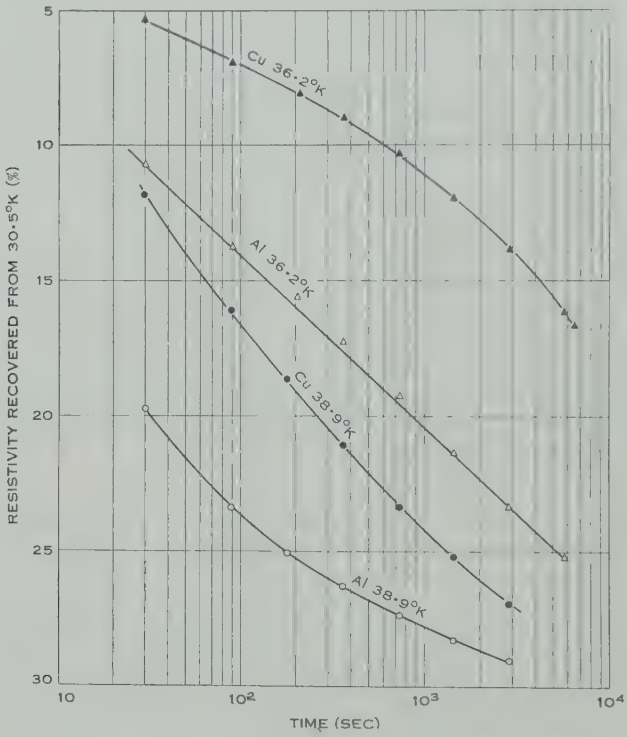


Fig. 5.—Isothermal pulse annealing of irradiated copper and aluminium at 36·2 and 38·9 °K.

lies in the fact that the method used to determine the activation energies makes the tacit assumption of a unique activation energy. The data from the isochronal anneals on the other hand suggest a continuous spectrum of energies, and it is possible that the analysis of the rate equation breaks down under this condition. At any rate, these experiments do point out the non-uniqueness of the activation energy in the range from 10 to 40 °K.

TABLE 3

AVERAGE NUMBER OF JUMPS OCCURRING BEFORE ANNIHILATION IN THE CASE OF COPPER AND ALUMINIUM*

T (°K)	ε (eV)	Copper n	Aluminium n
9·8	0·02	5×10^{4}	5×10^{4}
12·4	0·032	10^{2}	10^{2}
14·0	0·040	4	4×10^{-5}
15·1	0·065	5×10^{-6}	5×10^{-2}
16·2	0·080–0·068	1×10^{-9}	1×10^{-10}
17·2	0·095–0·088	2×10^{-12}	3×10^{-8}
37·4	0·180	2×10^{-10}	2×10^{-6}

* The values have been determined from the experimental activation energy and the rate equation; n is given by $t\nu e^{-\varepsilon/kT}$, where t is the time for annihilation, taken as 10^2 sec here, and ν is the atomic frequency factor, assumed to be 10^{13}.

V. ACKNOWLEDGMENTS

It is a pleasure for the authors to acknowledge frequent discussions with J. Diehl, G. Leibfried, and D. K. Holmes. They are also grateful for the assistance of J. K. Redman in making the measurements and for the encouragement of D. S. Billington.

VI. REFERENCES

COLTMAN, R. R., BLEWITT, T. H., and NOGGLE, T. S. (1957).—*Rev. Sci. Instrum.* **28**: 375.
CORBETT, J. W., and WALKER, R. M. (1958).—*Phys. Rev.* **110**: 767.
HOWE, J. T., COLTMAN, R. R., and BLEWITT, T. H. (1957).—Proc. 1957 Cryogenic Engineering Conference, p. 173.
HUNTINGTON, H. B. (1953).—*Phys. Rev.* **91**: 1092.

THE INSTABILITY OF PLASTIC FLOW OF METALS AT VERY LOW TEMPERATURES. II

By Z. S. Basinski*

[*Manuscript received September 25, 1959*]

Summary

Evidence is presented to show that the load drops observed in most metals at very low temperatures arise from thermal instability rather than mechanical instability of the lattice. Measurements of the temperature rise are described.

I. Introduction

Discontinuous load drops occurring during the deformation of metals at low temperatures have been reported by many workers, and appear to be quite a general phenomenon. For example, previous work on 24S aluminium alloy (Basinski 1957) which will be referred to as I, showed that the deformation at temperatures below about 10 °K takes place by a series of abrupt drops in load sometimes separated by regions of smooth plastic flow. This effect has been ascribed variously to the formation of mechanical twins (Blewitt, Coltman, and Redman 1955; Smith and Rutherford 1957), strain-induced martensitic transformation (Uzhik 1955), catastrophic breakdown of the dislocation barriers at the high stresses involved (Seeger 1957; Haasen 1958*a*, 1958*b*), and thermal instability of deformation resulting from the large temperature dependence of the flow stress coupled with the low specific heat (Wessel 1957; Basinski 1957). To date there seems to be little general agreement as to the mechanism underlying the appearance of these discontinuities. In cases where mechanical twins or strain-induced martensite can be formed, they will undoubtedly contribute to the discontinuities; however, the examination of published results, with the exception of the work of Blewitt, Coltman, and Redman on copper single crystals, shows that the strain could not be entirely accounted for by the observed twins. Where twin formation does not occur, the two possibilities left are the breakdown of dislocation barriers and adiabatic deformation.

In I it has been shown experimentally that the temperature of the specimen rises above the superconducting transition temperature of niobium (~9 °K) during a load drop. To account for the magnitude of the drop in load, however, on the adiabatic deformation model, the temperature would have to rise to 50–80 °K. In view of the fact that during this type of experiment the specimen was always immersed in liquid helium and that the heat losses might therefore be quite considerable, it seemed desirable to check experimentally whether the conclusion reached in I, of approximately adiabatic conditions during the load drop, actually held.

---

* Division of Pure Physics, National Research Council, Ottawa, Canada.

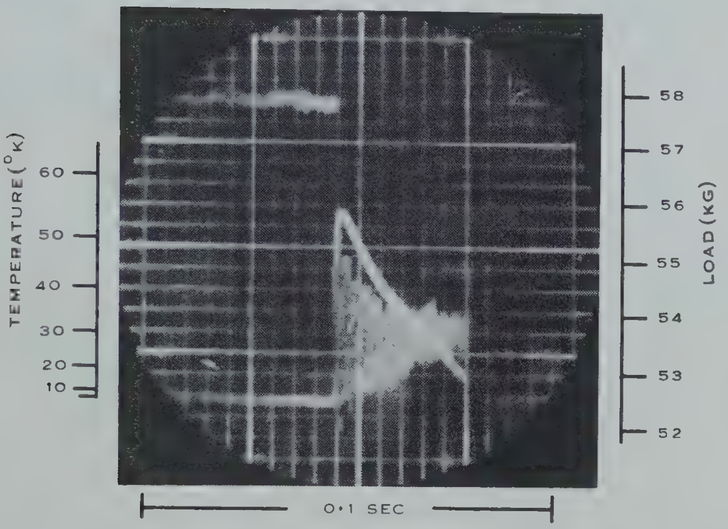


Fig. 1.—Oscilloscope trace showing a load drop and a temperature rise in deforming region produced by it.

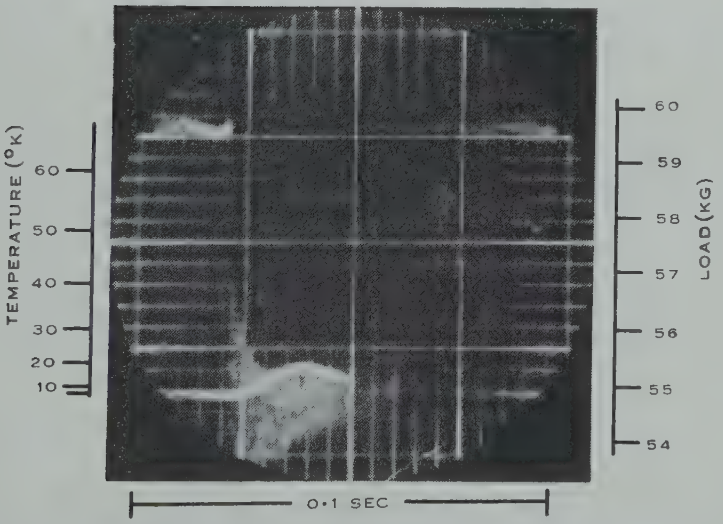


Fig. 2.—The temperature rise produced by the deformation not in the region of the thermocouple. Note about 5 msec delay between load drop and temperature rise.

II. EXPERIMENTAL METHOD

Specimens of 65S aluminium alloy were prepared by cold drawing to 1·2 mm diameter wire, annealing for 15 hr at 500 °C, and quenching in ice-brine mixture. The gauge length of the specimens was 5 cm.

The twisted end of uninsulated AuCo-AgAu thermocouple wire (0·003 in. in dia.) was inserted into a groove in the specimen cut with a razor blade parallel to the axis, and secured by gently squeezing the metal round it. The reference junction of the thermocouple was held at constant temperature by a helium bath surrounding the specimen. With this arrangement the thermocouples should indicate the temperature of a point on the specimen surface. The thermocouple was connected to one channel of a twin beam oscilloscope via a wide-band low-level amplifier with calibrated gain.

The output of the resistance strain gauge load cell was connected to a usual pen recording system, and via another wide-band amplifier to the second channel of the oscilloscope. The sensitivity of the load cell, as displayed on the oscilloscope, was calibrated by a simulated load of 4 kg produced by the zero suppression circuit (I).

In all cases the specimen was completely immersed in liquid helium and the deformation was carried out at a tensile strain rate of 10^{-3} per sec. The oscilloscope screen was filmed simultaneously with the recording of the normal load-elongation curve so that 1 : 1 correspondence could be established between the load drops recorded on the chart and on the oscilloscope.

The stress-strain curves obtained were very similar to those for 24S alloy, showing load drops increasing in frequency and magnitude during the course of deformation. The temperature traces during the load drops could be roughly classified into three groups. About 8–10 per cent. of the abrupt load drops were associated with a sharp temperature rise followed by a much slower decay, as illustrated by the traces shown in Figure 1. The smooth trace, which starts at a low level (left-hand side of photograph) then increases sharply and drops more slowly, represents the thermocouple output, the wider trace, which starts at a high level and drops sharply followed by rapid oscillations, represents the output of the load cell. The oscillations following the load drop are always observed, and most probably represent the vibrations in the tensile testing machine excited by the rapid decrease in load. The highest temperatures reached here varied between 45 and 60 °K.

About 20 per cent. of the load drops were accompanied by a much slower temperature rise. In many cases a delay of the order of 1–5 msec was observed between the beginning of the load drop and the beginning of the temperature rise. Such a trace is shown in Figure 2. There was wide variation in the top temperatures reached in this type of event, varying from barely detectable to 35–40 °K. No temperature change could be observed in about 70 per cent. of the load drops.

Obviously these three types of temperature change associated with the load drops are not fundamentally different, but the sharp rise in temperature corresponds to deformation in the region of the specimen where the thermocouple is situated. The rounded temperature peaks represent deformation occurring in regions near to the thermocouple, so that the heat conducted down the specimen is recorded. When no temperature change was observed during a load drop, the deformation most probably occurred in regions far removed from the thermocouple, so that the heat was dissipated before reaching it.

The top temperature T reached during a load drop under purely adiabatic conditions can be obtained from the relation

$$W\delta l \simeq v\int_{4\cdot 2}^{T} c_p \mathrm{d}T,$$

where W is the load on the specimen, δl the extension during a load drop, v the volume of the region in which catastrophic flow occurs, and c_p the specific heat per unit volume, which here is assumed to be given reasonably well by the Debye model with characteristic temperature of 400 °K. The value of v can be obtained from the relative frequency of the sharp temperature peaks, since the ratio of the volume in which deformation takes place to the total volume of the specimen should be approximately equal to the ratio of the number of sharp temperature peaks to the total number of load drops, and will be about 4×10^{-3} cm^3. Using the values of $W=58$ kg and $\delta l=7\cdot 5\times 10^{-3}$, corresponding to the load drop in Figure 1, obtained from the load-elongation curve, we get $T=61$ °K, which compares well with the measured value of 55 °K. The mean temperature in the deforming region is presumably somewhat higher than that indicated by the thermocouple, since this measures the temperature on the surface, where the cooling by the surrounding liquid helium would be largest.

The conditions during the deformation therefore are approximately adiabatic, as concluded in I. This is not surprising since the rate of heating of the order of $\sim 2\cdot 5\times 10^5$ °/sec is much faster than that of subsequent cooling, $\sim 1\cdot 5\times 10^3$ °/sec.

The magnitude of the load drop produced by a temperature rise to 55 °K will be given by

$$\delta W = W(\sigma_T/\sigma_{4\cdot 2}-1)+\delta W_{SH},$$

where δW_{SH} is the change of strength due to work hardening in the deforming region and $\sigma_T/\sigma_{4\cdot 2}$ is the ratio of the flow stresses at temperatures T and $4\cdot 2$ °K. δW_{SH} can be estimated from the load-elongation curve. The variation of flow stress with temperature was determined for 24S alloy in I; if we assume that the same temperature dependence holds also for 65S alloy, then the calculated value of δW is about 8 kg, whereas the measured value is $6\cdot 1$ kg. Adiabatic heating can thus account for the total load drop.

III. Discussion

In the present case, where neither twinning nor martensitic transformation takes place during the deformation, the only alternative explanation which has been suggested for the occurrence of the instability of plastic flow at very low temperatures other than the mechanism described here is that of Seeger (1957) and of Haasen (1958*a*, 1958*b*), in which the load drops are ascribed to the catastrophic breakdown of the dislocation barriers under the high stresses reached at these temperatures. In I it was observed that in 24S alloy the discontinuities during the deformation did not occur above about 12 °K. It would be surprising if a decrease in stress of only $\sim 2 \cdot 5$ per cent., which results from warming the specimen to 12 °K, could cause a complete change in the characteristics of the deformation. Furthermore, the stresses reached towards the end of the deformation at 12 °K, where no load drops are observed, were about double those in the early stages of deformation at $4 \cdot 2$ °K, where serrations were quite frequent. These arguments show that the explanation in terms of the high stresses reached is inadequate. On the other hand, the adiabatic deformation picture relies essentially on the small magnitude of the specific heat, and this changes by more than an order of magnitude between 4 and 12 °K.

The objection has been raised (Haasen 1958*b*) that " considerable localized plastic flow must have occurred already *before* the temperature will rise ". At $4 \cdot 2$ °K the energy liberated during ordinary slip line formation may well be sufficient to produce heat pulses strong enough to start the catastrophic flow.

Obviously, in cases where mechanical twins or strain-induced martensite form, they can serve as nucleating agents, thus enabling adiabatic deformation to occur at higher temperatures. This may be one reason why discontinuous load drops in iron can be observed at temperatures as high as 78 °K (Smith and Rutherford 1957), while the twins themselves can account for only a fraction of the observed deformation. Another factor which may be important in the case of iron is the much steeper temperature dependence of the flow stress (Basinski and Christian 1960).

Since the nucleation of load drops is the most difficult part of the process, we may expect that the critical temperature for the drops to occur will be much lower than the maximum temperature reached once the process has started.

IV. ACKNOWLEDGMENTS

The author is grateful to Dr. S. B. Woods for suggesting the method of attaching the thermocouple to the specimen and to Dr. J. S. Dugdale for comments on the manuscript.

V. REFERENCES

BASINSKI, Z. S. (1957).—*Proc. Roy. Soc.* A **240** : 229.

BASINSKI, Z. S., and CHRISTIAN, J. W. (1960).—*A. J. Phys.* **13** : 299.

BLEWITT, T. H., COLTMAN, R. R., and REDMAN, J. K. (1955).—" Defects in Crystalline Solids." (Phys. Soc. : London.)

HAASEN, P. (1958*a*).—*Phil. Mag.* **3** : 384.

HAASEN, P. (1958*b*).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **212** : 42.

SEEGER, A. (1957).—" Dislocations and Mechanical Properties of Crystals." p. 206. (Wiley : New York.)

SMITH, R. L., and RUTHERFORD, J. L. (1957).—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **209** : 857.

UZHIK, G. V. (1955).—*Izvest. Akad. Nauk S.S.S.R. Otdel. Tek. Nauk* **1** : 57.

WESSEL, E. T. (1957).—*Trans. Amer. Soc. Metals* **49** : 149.

A LOW TEMPERATURE YIELD INSTABILITY IN IRON

By J. B. Lean*

[*Manuscript received September* 25, 1959]

Summary

Polycrystalline specimens of a $0 \cdot 07\%$ C steel were tested in tension at liquid air temperatures. Under particular conditions of temperature and strain rate, following the upper yield point, the recorded stress descended below the lower yield stress then returned to the lower yield stress quasi-elastically thus forming a serration on the stress-strain curve. The magnitude of the serration depended on the difference between the upper and lower yield stresses, on the rigidity of the testing machine employed, and on the applied strain rate.

In the discussion the mechanical stability of the tensile test during the yield point load drop is examined. The effect of adiabatic temperature increase in specimens deformed at low temperatures in boiling liquids is also discussed.

I. Introduction

Recent investigations of the mechanism of deformation of metals at low temperatures have revealed regions of unstable plastic flow. The instabilities occur in a number of metals at 4 °K, including copper single crystals (Blewitt, Coltman, and Redman 1957) and aluminium alloys (Basinski 1957) ; below 45 °K austenitic stainless steels exhibit serrated stress-strain curves (Uzhik 1955) ; and a number of authors have noted discontinuous yield in iron single crystals below 100 °K, e.g. Allen, Hopkins, and McLennan (1956).

The instabilities have been attributed to crystallographic changes, twinning, and the formation of martensite, or to adiabatic temperature increase in the deforming specimen. On the other hand Kochendörfer and Wink (1957), discussing effects due to the hardness of the testing machine at the yield point in steels, have shown theoretically that damped vibrations at the yield point drop in load can occur, and that these vibrations are due to the mechanical instability of the tensile test.

In the present communication, instabilities, associated with the yield point load drop, noted in low temperature experiments, will be discussed using a model based on that developed by Kochendörfer and Wink but differing from their model in introducing a more precise non-linear damping term.

The instability caused by adiabatic increase of temperature within a moving Lüders band in a steel specimen tested in a low temperature boiling liquid will also be discussed.

II. Experimental Procedure

The present investigation arose from observations of yield instability made during a study of the dispersion, in the tensile test, of the brittle fracture transition temperature of low carbon steels (Lean and Plateau 1958).

---

* Broken Hill Pty. Co. Ltd., Central Research Laboratories, Shortland, N.S.W.

The tensile testing machine, constructed for the investigation, was a rigid assembly within which the load, applied through a screw, was transmitted to the specimen, immersed in a constant temperature bath, by a rod in series either with a strain gauge load cell (hard machine) or with a proving ring (soft machine). The mass of the 1·8 cm diameter stainless steel rod was 600 g. Including the mass of the proving ring, the inertial mass in the soft machine was 1·5 kg. The extension of the proving ring under load was 0·25 cm for a load of 1000 kg. Displacement of the proving ring moved an optical tripod, enabling the load-elongation curve to be photographically recorded.

The test specimens used were 2 cm long and had a diameter of 0·25 cm. The applied strain rates varied from $1 \cdot 4 \times 10^{-5}\ \mathrm{sec}^{-1}$ to $3 \cdot 8 \times 10^{-1}\ \mathrm{sec}^{-1}$ in factors of 30. Since, depending on the strain rate, brittle fracture occurred at, or slightly below liquid air temperatures, the lowest temperature used was −210 °C. Tests were carried out on a steel analysing 0·07% C, 0·43% Mn, with a mean grain size of 1450 grains/mm^2.

III. Experimental Results

In tests carried out in the hard tensile testing machine at a strain rate of $4 \cdot 2 \times 10^{-4}\ \mathrm{sec}^{-1}$ no abnormal effects were noted, but in tests carried out in the soft machine, at temperatures immediately above those at which brittle fracture occurred, a region of instability at the lower yield point was observed. Typical stress-strain curves obtained on the soft machine are shown in Figure 1. The instability is evidenced as a serration on curves B, C, and D and by the failure recorded in curve E.

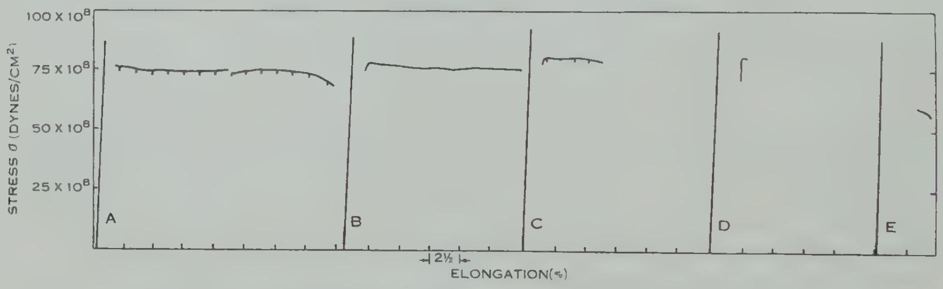


Fig. 1.—Yield instability in low temperature stress-strain curves.
Curve A, $T = 90\ °\mathrm{K}$; $\dot{\varepsilon}_i = 4 \cdot 2 \times 10^{-4}\ \mathrm{sec}^{-1}$; timing trace at 32·5 sec intervals.
Curve B, $T = 90\ °\mathrm{K}$; $\dot{\varepsilon}_i = 1 \cdot 27 \times 10^{-2}\ \mathrm{sec}^{-1}$.
Curve C, $T = 77\ °\mathrm{K}$; $\dot{\varepsilon}_i = 4 \cdot 2 \times 10^{-4}\ \mathrm{sec}^{-1}$; timing trace at 32·5 sec intervals.
Curve D, $T = 86\ °\mathrm{K}$; $\dot{\varepsilon}_i = 1 \cdot 27 \times 10^{-2}\ \mathrm{sec}^{-1}$.
Curve E, $T = 102\ °\mathrm{K}$; $\dot{\varepsilon}_i = 3 \cdot 8 \times 10^{-1}\ \mathrm{sec}^{-1}$.
(T, Bath temperature; $\dot{\varepsilon}_i$, impressed strain rate.)

On curves B, C, and D a hiatus, occurring after the upper yield point, is followed by a stress minimum; the stress then increasing to the lower yield point to form the serration. Due to the photographic recording method employed the extremely rapid deformations occurring after the upper yield point have not been registered. Check studies made with intense illumination and rapid film showed the missing portion of the curves to be a nearly straight line connecting the upper yield point to the remainder of the curve.

The serration was not a spurious effect introduced by the recording mechanism. The test could be halted during the quasi-elastic stress rise following the stress minimum and the optical tripod was not dislodged. The regular timing marks on curves A and C are not serrations.

Comparison of curve A with B demonstrates the sensitivity of the serration to increase in the impressed strain rate ; comparison of curve A with C, and B with D, illustrates the sensitivity of the serration to decrease in testing temperature. Changes in temperature and strain rate affect the difference between the upper and lower yield stresses and further experiments were required to distinguish the significant variables.

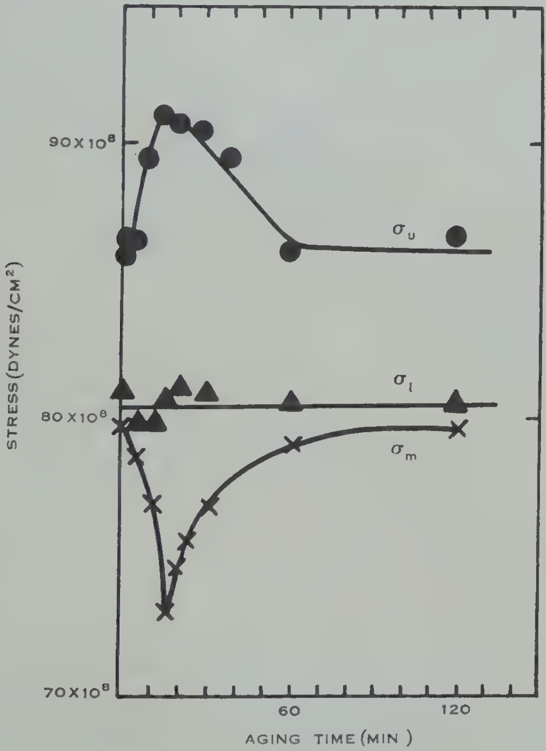


Fig. 2.—Dependence of σ_m, the minimum stress in the serration, on the difference between the yield stresses σ_u and σ_l. Samples aged in boiling water, tested at 90 °K ; $\dot{\varepsilon}_i = 1\cdot27\times10^{-2}$ sec^{-1}.

Specimens were artificially aged in boiling water and after varying times were quenched into liquid nitrogen and tensile-tested at -183 °C, in liquid oxygen, at a strain rate of $1\cdot27\times10^{-2}$ sec^{-1}. The results are plotted in Figure 2. The minimum stress registered in the serration, σ_m, is dependent on the difference between the upper yield stress, σ_u, and the lower yield stress, σ_l. The lower yield stress was not significantly affected by the aging treatment.

Similar experiments were carried out over a range of temperatures at two strain rates and the results are presented in Figure 3. Within the scatter the results at one strain rate all lie on a curve expressing a relationship between the difference between the upper and lower yield stresses ($\sigma_u - \sigma_l$) and the difference between the lower yield stress and the minimum stress ($\sigma_l - \sigma_m$). To a critical difference in the yield stresses no serration is observed, thereafter the relationship is linear with unit slope, departing from linearity only at high values of the stress differences. At an impressed strain rate of $4 \cdot 2 \times 10^{-4}$ sec^{-1} the critical difference in yield stress is $10 \times 10^8 \pm 1 \times 10^8$ dynes cm^{-2} and at $1 \cdot 27 \times 10^{-2}$ sec^{-1} it is $5 \times 10^8 \pm 1 \times 10^8$ dynes cm^{-2}. In the range of temperatures explored, within the accuracy of the measurements, the sensitivity of the serration to change in testing temperature is due to the effect of temperature on the difference between the yield stresses. At liquid air temperatures, at the lowest strain rate, $1 \cdot 4 \times 10^{-5}$ sec^{-1}, no serrations were observed; the difference in the yield stresses being insufficient to induce the effect. At the highest strain rate, $3 \cdot 8 \times 10^{-1}$ sec^{-1}, brittle fracture occurred in this temperature range.

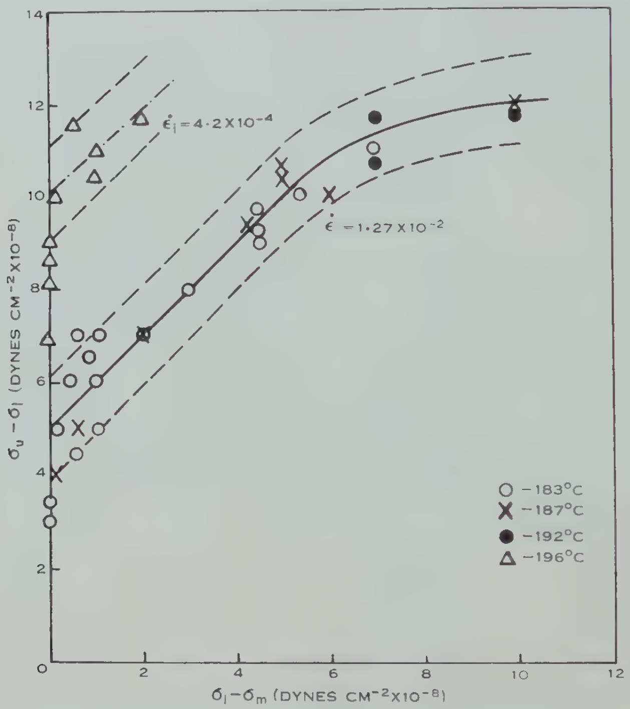


Fig. 3.—Relationship between the yield stresses, σ_u and σ_l, and the minimum stress recorded in the serration, σ_m, at strain rates at $4 \cdot 2 \times 10^{-4}$ sec^{-1} and $1 \cdot 27 \times 10^{-2}$ sec^{-1} at temperatures between 77 and 90 °K.

A number of broken test-pieces were aged and macro-etched in Fry's reagent. In cases where the serration did not appear or was of small amplitude the yield point extension invaded the working length of the specimen (curves type A and B, Fig. 1). As the testing conditions increased in severity, deformation became localized. In tests arrested after the serration during the quasi-elastic stress rise, intense plastic deformation, evidenced by a constriction, was confined to a limited length of the specimen.

In curves of type C, Figure 1, the specimens were found to break either at the minimum section or in a brittle manner in metal adjacent to the strain-hardened zone (e.g. curve D).

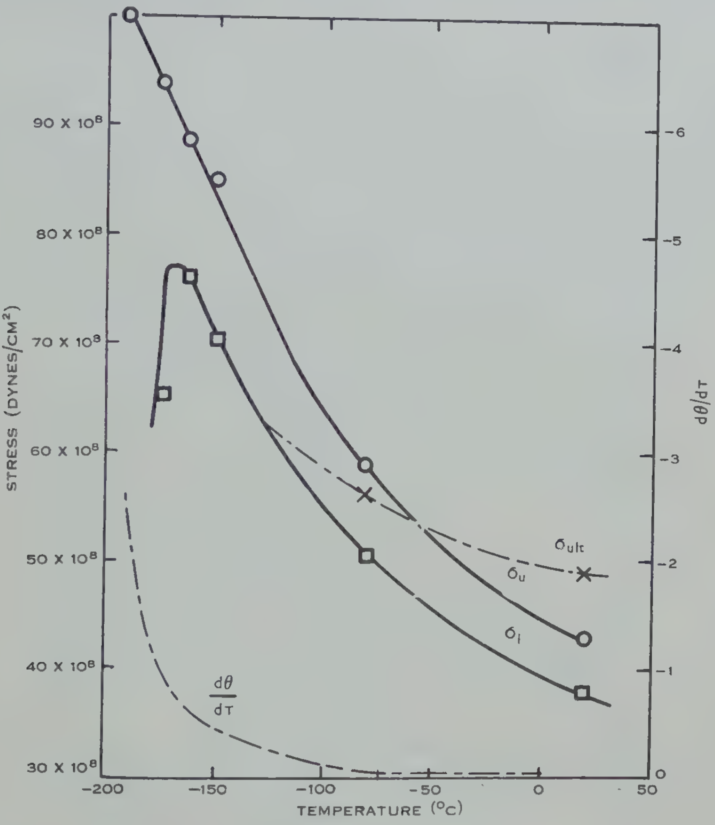


Fig. 4.—Temperature variation of the upper yield stress σ_u; the apparent lower yield stress σ_l; the ultimate stress σ_{ult}; and of $d\theta/dT$, the theoretical maximum rate of increase of adiabatic temperature rise θ, with testing temperature T, at a strain rate of $3 \cdot 8 \times 10^{-1}$ sec^{-1}.

Some twins, Neumann bands, were visible within isolated grains in polished and etched sections of the specimens, and in particular at the minimum section, adjacent to the fracture surface, but there appeared to be no significant increase in the number of twins in specimens which displayed the serration effect when these were compared with specimens which exhibited normal stress-strain curves.

Specimen E, tested at high strain rate, $3 \cdot 8 \times 10^{-1}$ sec^{-1}, and slightly higher temperature, broke in a ductile manner with 68 per cent. reduction of area in a restricted deformation band. The variation of the yield and ultimate stresses over a wider temperature interval at this strain rate is shown in Figure 4. Also shown in the figure is the variation of the temperature rate of adiabatic temperature rise in the specimen, calculated from the work done on the specimen during the yield extension and from values quoted by Kenyon (1948) for the specific heat of iron.

The figure suggests that adiabatic temperature rise in the specimen may be the cause of the yield instability. Due to the proximity of the temperature at which serrations were observed to the ductility transition temperature, attempts to measure adiabatic temperature increases failed. The specimens broke in a brittle manner at the points of attachment of the thermocouples. Without intimate contact of the thermocouples and the specimen only the bath temperature would have been measured.

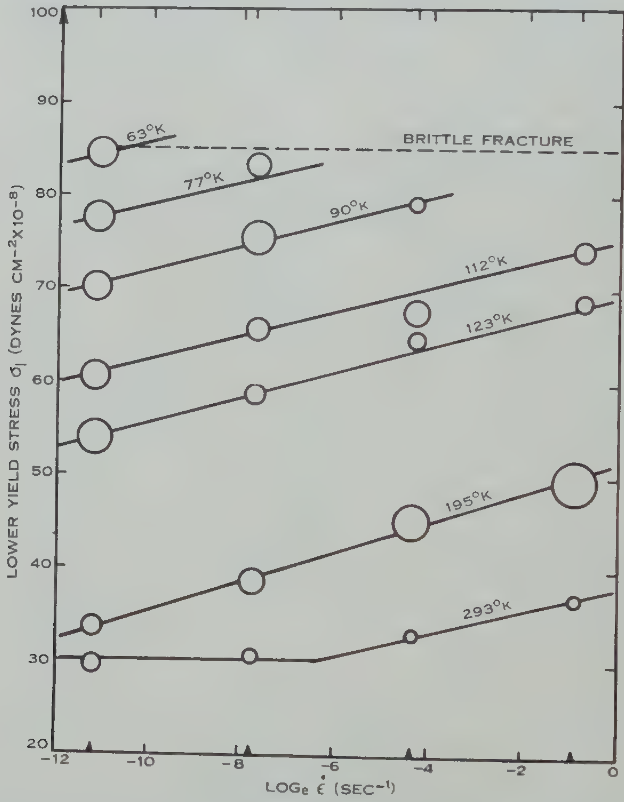


Fig. 5.—Variation of the lower yield stress with strain rate at temperatures between 63 and 293 °K. (The radius of each circle represents the 95% confidence limits of the determination of the mean stress.) $d\sigma_l/d\ln\dot{\varepsilon}=1\cdot3\times10^8$ dynes/cm^2.

From the arguments advanced by Kochendörfer and Wink (loc. cit.) the serration may have been caused by the mechanical instability of the soft tensile test employed. In order to obtain a precise estimate of the damping constant of the system (see Discussion, Section IV) the effect of the strain rate on the flow stress of the specimens was examined and the results are presented in Figure 5.

IV. Discussion

(*a*) *Mechanical Vibrations caused by the Yield Point Drop in Load*

A typical arrangement of the deforming members in a tensile test is shown in Figure 6.

The overall increase, ΔL cm in length L of the system, in time t' sec after the test commences at a nominal impressed strain rate of $\dot{\varepsilon}_i$ sec^{-1} on a test-piece of length l, is $\dot{\varepsilon}_i l t'$ and the changes in length of the individual members of the system sum to this value :

$$\Delta L = \Delta l + \Delta s, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

$$\dot{\varepsilon}_i l t' = (\sigma_T/E_T + \varepsilon_p + \sigma_S/E_S) l, \quad \ldots\ldots\ldots \quad (2)$$

where s is the length of the spring ; σ_T dynes cm^{-2} is the elastic stress in the test-piece, σ_S is the elastic stress in the spring, ε_p is the plastic strain in the test piece ; E_T dynes cm^{-2} is the Young's modulus of the material tested ; and E_S, a modulus relating the equivalent strain in the spring to the stress rise in the specimen, is defined by

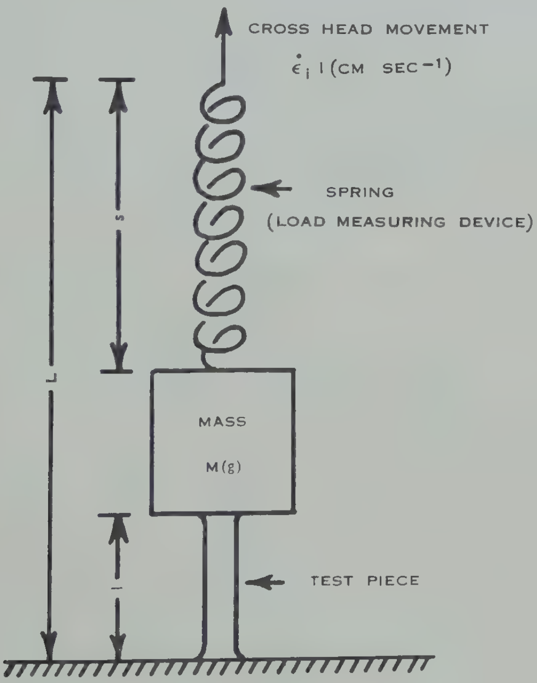


Fig. 6.—Schema of a typical tensile test.

$$E_S = (\mathrm{d}W/\mathrm{d}s) l/a,$$

where $\mathrm{d}W/\mathrm{d}s$ is the spring constant of the load-measuring device in dynes cm^{-1}.

If it is assumed that all the inertial mass in the system is concentrated at a point between the spring and the specimen, then, from (2), the elastic stresses are related to ε', the total strain in the specimen (which determines the motion of the mass M g) by

$$\varepsilon' = \sigma_T/E_T + \varepsilon_p = \dot{\varepsilon}_i t' - \sigma_S/E_S, \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

that is,

$$\sigma_S = E_S(\dot{\varepsilon}_i t' - \varepsilon'), \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

and

$$\sigma_T = E_T(\varepsilon' - \varepsilon_p). \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

The forces acting on the mass M during the load drop are

$$Ml\ddot{\varepsilon}'/a = \sigma_S - (\sigma_T + \sigma_D), \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

where σ_D is the strain rate dependent part of the stresses opposing motion of the mass. While, from (4), σ_S is known, $\sigma_T + \sigma_D$ is not measured and must be inferred from other evidence.

Kochendörfer and Wink (1957) derived an estimate for the damping forces in a tensile test system by assuming:

1. that the instantaneous stress, $\sigma_T + \sigma_D$, during the yield jog is equal to the lower yield stress of the material measured at an impressed strain rate equal to the instantaneous plastic strain rate $\dot{\varepsilon}_p$, during the yield extension, where $\dot{\varepsilon}'$ is constant,
2. that the relationship between strain rate and yield stress is linear.

Their equation (4), in the system used here, is

$$(\sigma_T + \sigma_D) = \sigma_{l0} + K_d \dot{\varepsilon}_p \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

$$= \sigma_l + K_d(\dot{\varepsilon}_p - \dot{\varepsilon}_i), \quad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

where σ_{l0} is a static lower yield stress, σ_l is the lower yield stress at the impressed strain rate, and K_d is a damping constant. The assumptions in this relationship require careful examination.

By differentiating (2) with respect to time, it can be shown that opportunity is available for inhomogeneous plastic deformation to occur while the elastic stresses in the system are increasing, provided the overall plastic strain rate is lower than the impressed strain rate. We have obtained experimental evidence which we have interpreted as indicating that Lüders band formation commences before the upper yield point (Lean and Plateau 1959). The behaviour of steel during the yield extension should then be a reliable guide to behaviour during the yield drop, since the mechanism of deformation in the two regions is identical. In the absence of information on strain hardening in steel at low and inhomogeneous deformations, Kochendörfer and Wink's first assumption appears reasonable.

Expressions based on this assumption should only be applied to increasing strain rates or to systems in which the stress-strain curve of the material tested does not display marked strain hardening (e.g. steel at low temperatures), otherwise the damping stress, should the strain rate ultimately decrease, could be seriously underestimated, since increase in the lower yield stress with increase in strain rate is accompanied by increase in the yield strain (Lean, Plateau, and Crussard 1958).

Figure 5 indicates that, although it is possible to find a static lower yield stress at room temperature, in the range of strain rates employed no static stress is found at lower temperatures. The room temperature static yield is exceptional and is probably due to the diffusion of atmospheres of interstitial atoms at rates comparable with the speed of displacement of dislocations in slowly strained specimens or to the partial re-formation of atmospheres about temporarily blocked dislocations. These processes, requiring thermal activation, could be effective only at extremely slow rates of deformation at lower temperatures.

From Figure 5 the lower yield stress varies linearly with the logarithm of the strain rate. Therefore we write instead of (**8**), for positive values of $\dot{\varepsilon}_p$,

$$\sigma_T+\sigma_D=\sigma_l+K_d \ln (\dot{\varepsilon}_p/\dot{\varepsilon}_i), \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{9})$$

where the mean value of K_d is $1{\cdot}3\times10^8$ dynes cm^{-2} for our specimens.

For a soft tensile test we assume that $\dot{\varepsilon}'$ is approximately equal to $\dot{\varepsilon}_P$. A slight correction is required for strain rates approaching the strain rate in elastic extension. During elastic extension $\sigma_S=\sigma_T=\sigma$ the elastic stress in the system and from (**2**)

$$\dot{\varepsilon}'=\dot{\varepsilon}_p+\dot{\sigma}/E_T=\dot{\varepsilon}_i-\dot{\sigma}/E_S, \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{10})$$

that is, as the strain rate approaches the elastic rate,

$$\frac{\dot{\varepsilon}_p}{\dot{\varepsilon}_i}=\frac{\dot{\varepsilon}'}{\dot{\varepsilon}_i}\left(1+\frac{E_S}{E_T}\right)-\frac{E_S}{E_T}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{11})$$

due to its logarithmic dependence in (**9**) the damping stress is not sensitive at higher strain rates to the small errors this expression introduces into the $\dot{\varepsilon}_p/\dot{\varepsilon}_i$ term. Equation (**6**) then becomes

$$\frac{lM\ddot{\varepsilon}'}{a}=E_S(\dot{\varepsilon}_i t'-\varepsilon')-\left[\sigma_l+K_d \ln \left\{\frac{\dot{\varepsilon}'}{\dot{\varepsilon}_i}\left(1+\frac{E_S}{E_T}\right)-\frac{E_S}{E_T}\right\}\right]. \quad \ldots\ldots \quad (\mathbf{12})$$

At σ_u, the upper yield stress,

$$\left.\begin{aligned} \dot{\varepsilon}'&=\dot{\varepsilon}_i,\\ t'&=\frac{\sigma_u}{\dot{\varepsilon}_i}\left(\frac{1}{E_S}+\frac{1}{E_T}\right),\\ \varepsilon'&=\sigma_l/E_T+\varepsilon_p, \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{13})$$

and taking $t=0$, $\varepsilon=0$, assuming $\varepsilon_p=(\sigma_u-\sigma_l)/E_T$, substituting for t' and ε', and rearranging, we obtain

$$\frac{lM\ddot{\varepsilon}}{a}+K_d \ln \left[\frac{\dot{\varepsilon}}{\dot{\varepsilon}_i}\left(1+\frac{E_S}{E_T}\right)-\frac{E_S}{E_T}\right]+E_S\varepsilon=E_S\dot{\varepsilon}_i t+\sigma_u-\sigma_l. \quad \ldots\ldots \quad (\mathbf{14})$$

Differentiating (**14**) and substituting $\ddot{\varepsilon}\,\mathrm{d}\ddot{\varepsilon}/\mathrm{d}\dot{\varepsilon}=\dddot{\varepsilon}$ an expression,

$$\ddot{\varepsilon}=\frac{-E_S(\dot{\varepsilon}-\dot{\varepsilon}_i)\left(\dot{\varepsilon}-\dot{\varepsilon}_i\left[\frac{E_S}{E_S+E_T}\right]\right)}{K_d+\frac{\mathrm{d}\ddot{\varepsilon}}{\mathrm{d}\dot{\varepsilon}}\,\frac{lM}{a}\left(\dot{\varepsilon}-\dot{\varepsilon}_i\left[\frac{E_S}{E_S+E_T}\right]\right)}, \quad \ldots\ldots\ldots\ldots \quad (\mathbf{15})$$

relating the strain rate to the strain acceleration is derived (McLachlan 1956). The general form of this relationship is shown schematically in Figure 7.

In interpreting this figure the following criteria apply.

(*a*) When $\dot{\varepsilon} > \dot{\varepsilon}_i$ the elastic stress in the spring, σ_S, is decreasing while the plastic strain in the test-piece is increasing.

(*b*) When $\dot{\varepsilon} = \dot{\varepsilon}_i$, by substitution in (**14**)

$$\sigma_S - \sigma_l = \ddot{\varepsilon} Ml/a. \qquad \textbf{(16)}$$

(i) The initial condition at the upper yield point is

$$\ddot{\varepsilon} = (\sigma_u - \sigma_l)a/Ml. \qquad \textbf{(17)}$$

(ii) A critical initial condition can be found such that $\ddot{\varepsilon} < 0$ when $\dot{\varepsilon}$ returns to $\dot{\varepsilon}_i$ at the end of the load drop. When $\ddot{\varepsilon} < 0$,

$$\sigma_m = \sigma_l + \ddot{\varepsilon} Ml/a, \qquad \textbf{(18)}$$

where σ_m is the minimum stress registered in a vibration associated with the load drop.

From the form of Figure 7, σ_m will decrease very nearly linearly with the increase of $\sigma_u - \sigma_l$ once the critical $\sigma_u - \sigma_l$ has been exceeded.

(iii) From (**15**) the critical initial condition will depend on the impressed strain rate (cf. Fig. 3).

(*c*) When $\dot{\varepsilon} < \dot{\varepsilon}_i$, the stress increases as the strain increases, and, when the critical initial conditions are exceeded and $\dot{\varepsilon}$ approaches $\dot{\varepsilon}_i E_S/(E_S + E_T)$ at the end of the load drop, the stress rise becomes quasi-elastic (cf. Fig. 1, curve D).

Making a number of simplifications, an analytical expression for the critical stress difference can be estimated from (**14**). Initially changes in t and ε can be neglected; then the decrease in acceleration will be due to increase in the damping stress σ_D due to increase in strain rate. At the maximum strain rate for a given stress difference, $\ddot{\varepsilon} = 0$ and

$$\dot{\varepsilon} = \{\dot{\varepsilon}_i/(E_S + E_T)\}[E_S + E_T \exp\{(\sigma_u - \sigma_l)/K_d\}]. \qquad \textbf{(19)}$$

Seeking a limit for this relationship, a linear approximation, valid over a very small change in strain rate, is substituted for (**14**)

$$\ddot{\varepsilon} + \left[\frac{K_d a(E_S + E_T)}{lM\dot{\varepsilon}_i E_T \exp\{(\sigma_u - \sigma_l)/K_d\}}\right]\dot{\varepsilon} + \left[\frac{E_S a}{lM}\right]\varepsilon$$
$$= \frac{K_d a}{lM}\left[1 + \frac{E_S}{E_T \exp\{(\sigma_u - \sigma_l)/K_d\}}\right]. \qquad \textbf{(20)}$$

In $\ddot{\varepsilon} + 2b\dot{\varepsilon} + c^2\varepsilon = \text{constant}$, the limit between the regions where periodic and aperiodic solutions apply is given by $b = c$.

The considerations leading to equation (**19**) also apply during deceleration when the periodic solution no longer operates (e.g. see Fig. 8), hence the critical stress difference is twice that estimated by the criterion $b = c$ and is given by

$$\sigma_u - \sigma_l \gtrsim 2K_d \ln\left\{\frac{K_d(E_S + E_T)}{2\dot{\varepsilon}_i E_T}\sqrt{\left(\frac{a}{MlE_S}\right)}\right\}. \qquad \textbf{(21)}$$

The expression developed in (**12**) predicts the existence of a highly damped vibration associated with the yield point drop in load which resembles that found in our experiments; however, substitution of values in (**21**) reveals that the predicted critical difference in yield stresses is much higher than that found in our experiments.

In Figures 8 and 9, curves I, corresponding to the critical initial conditions required by the theory, and the theoretical curves II, appropriate to the experimental critical initial conditions, are shown. Comparison of these figures reveals that the serration is only observed when strain rates of the order of, or greater than, 1 sec^{-1} can occur during the yield drop.

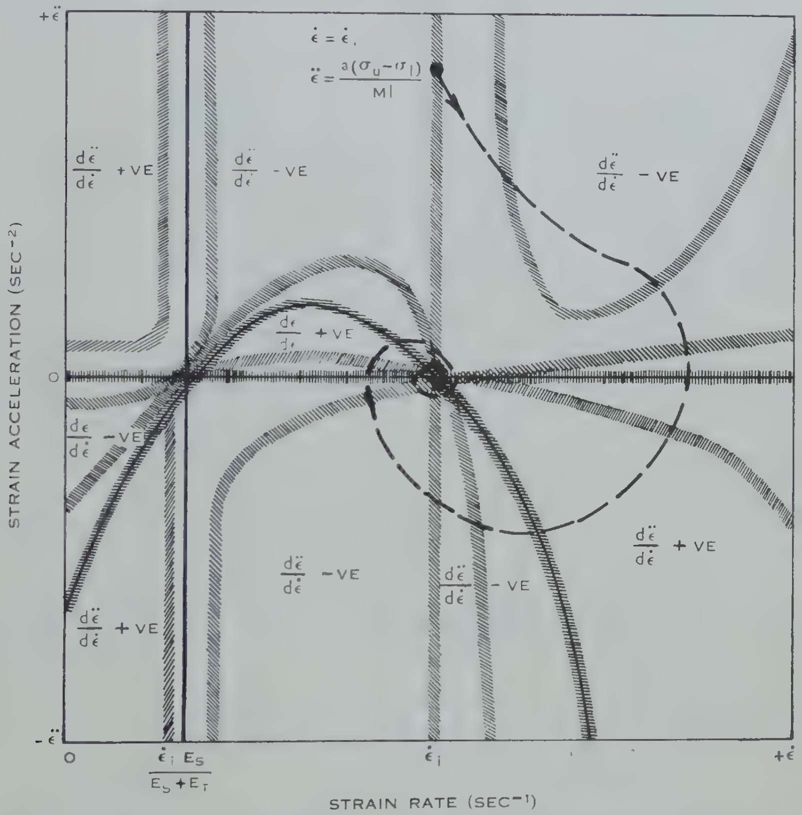

$+\ddot{\epsilon}$
$\dot{\epsilon} = \dot{\epsilon}_i$
$\ddot{\epsilon} = \frac{a(\sigma_u - \sigma_l)}{MI}$
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ +VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ -VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ -VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ +VE
O
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ -VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ +VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ -VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ -VE
$\frac{d\ddot{\epsilon}}{d\dot{\epsilon}}$ +VE
$-\ddot{\epsilon}$
STRAIN ACCELERATION (SEC^{-2})
O
$\frac{\dot{\epsilon}_i E_S}{E_S + E_T}$
$\dot{\epsilon}_i$
$+\dot{\epsilon}$
STRAIN RATE (SEC^{-1})

Fig. 7.—Schema of the relationship between the strain rate and strain acceleration (equation (**15**)) during the load drop in a tensile test on steel. The sketch is not to scale and the second vibration shown on the curve would normally not be observable.

The lack of agreement between the theory and the experiment could be due to a number of factors.

(i) *Overestimation of Strain Hardening in the Theory.*—If strain hardening to the lower yield stress is not established in the deformed regions of the specimen at the upper yield point, the initial acceleration will be underestimated and the damping stress overestimated. The effect of strain hardening cannot be evaluated but any increase in strain rate due to underhardening should be rapidly decelerated when gross plastic deformation commences.

(ii) *Strain-induced Crystallographic Changes.*—Twinning is strain rate and temperature dependent and could cause a load drop at a critical strain rate. However, the micrographic evidence that a few thin twins are formed in isolated grains does not suggest the existence of a sufficient volume of twinned material to introduce a load drop great enough to satisfy the difference between the experimental and theoretical results.

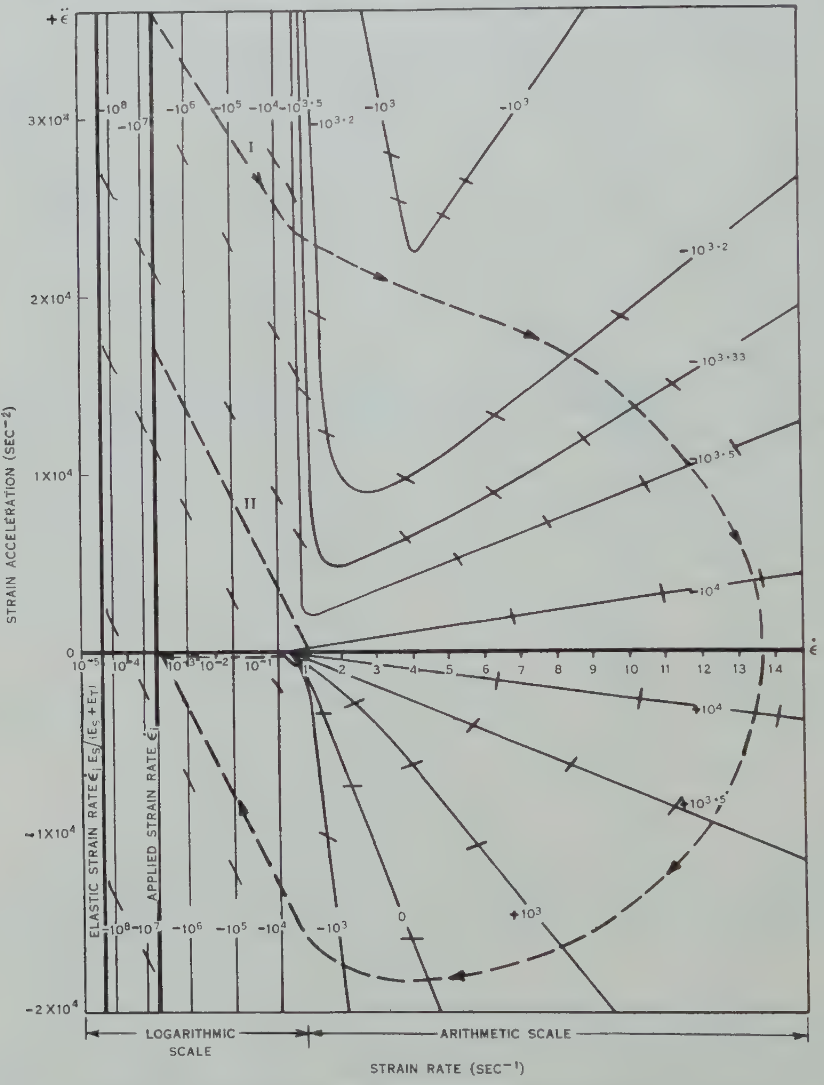


Fig. 8.—Relationship between strain rate and strain acceleration for tensile tests in the soft machine at an impressed strain rate of $4\cdot2\times10^{-4}$ sec^{-1}. Curve I is the theoretical curve for the critical difference of yield stresses, c. $21\cdot5\times10^{8}$ dynes cm^{-2}, beyond which a serration should occur. Curve II is a theoretical curve for the observed critical difference in yield stresses, 10×10^{8} dynes cm^{-2}. The figure is calculated for $E_T=1\cdot96\times10^{12}$ dynes cm^{-2}, $E_S=1\cdot57\times10^{11}$ dynes cm^{-2}, $K_d=1\cdot3\times10^{8}$ dynes cm^{-2}, $M=1500$ g, $l=2$ cm, $a=0\cdot05$ cm^{2}.

(iii) *Adiabatic Temperature Increase in the Deforming Specimen.*—From Figure 5, between 77 and 90 °K

$$\mathrm{d}\sigma_l/\mathrm{d}T = -7\times10^7 \text{ dynes cm}^{-2}\text{ degC}^{-1}. \qquad \ldots\ldots \quad (22)$$

A reversible temperature increase, at strain rates of 1 sec^{-1}, of 5 to 8 °C would reconcile the theoretical and observed critical stresses. From

$$\frac{J\sigma_l\varepsilon_y}{\rho} = \int_T^{T+\theta} C_p\mathrm{d}T, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (23)$$

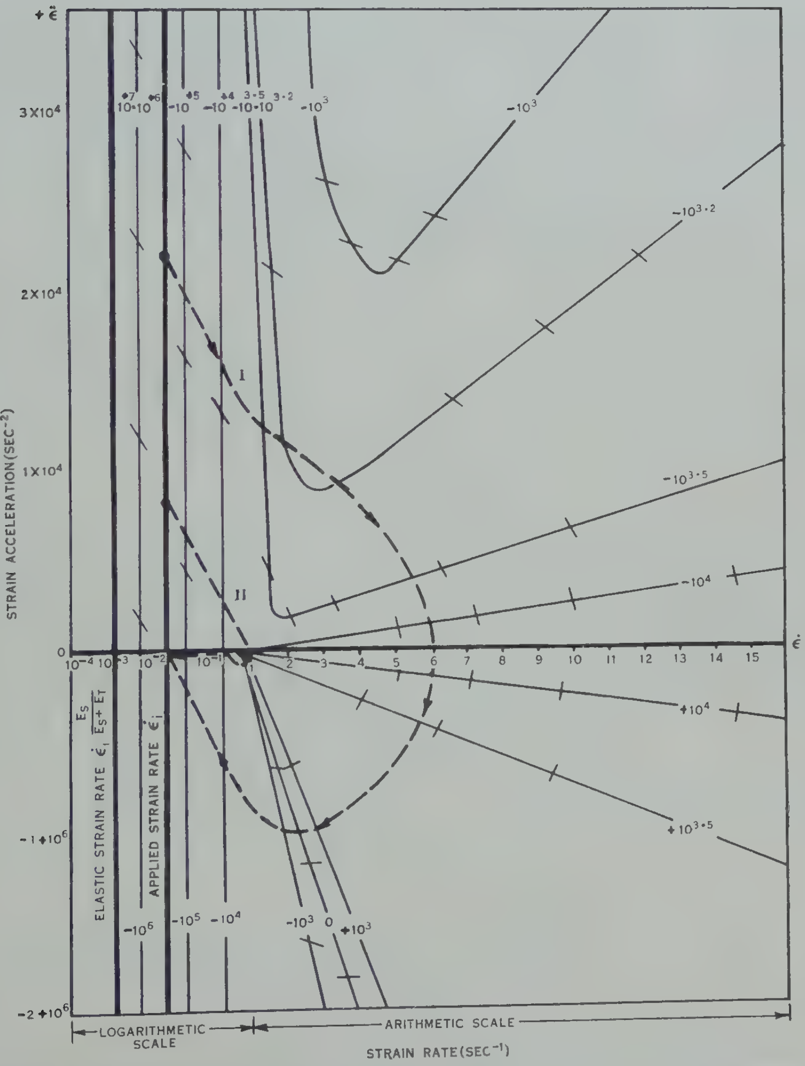


Fig. 9.—Relationship between strain rate and strain acceleration during tensile tests in the soft machine at an impressed strain rate of $1\cdot27\times10^{-2}$ sec^{-1} (refer Fig. 8).

where $\rho=7\cdot 8$ g cm^{-3}; C_p, for $75\ ^\circ\mathrm{K}<T<175\ ^\circ\mathrm{K}$, from Kenyon (1948) is approximately $5\times 10^{-4}T$ cal g^{-1} degC^{-1}; J is $2\cdot 39\times 10^{-8}$ calories erg^{-1}: $\sigma_l\approx 80\times 10^8$ dynes cm^{-2}; and ε_y, the yield strain, is $0\cdot 1$, temperature increases, θ, of 50 °C in the deformed specimen can occur if no heat is lost by conduction or to the bath. Adiabatic temperature rise appears to be the most probable mechanism of introducing the load drop.

(b) *Adiabatic Temperature Increase in Steel Deformed at Low Temperatures*

A qualitative estimate is sought of the relationship between the applied strain rate and the adiabatic temperature increase in specimens deforming inhomogeneously in low temperature boiling liquids. It is assumed that active deformation is confined within a thin plane moving along the test piece. Steel in the specimen in front of the plane has not been plastically deformed: steel behind the plane, in the Lüders band, has acquired the yield strain ε_y.

This model of Lüders band structure is greatly simplified and at low temperatures creep behind the working face of the band is undoubtedly important (Fisher and Rogers 1956). A wider zone of deformation will tend to assist retention of heat in the deformed region (Basinski 1957).

In the model the front of the band is a planar heat source and

$$q=J\sigma_l\dot{\varepsilon}_i l, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad \mathbf{(24)}$$

where q is the intensity of the source in calories cm^{-2} sec^{-1}. The source moves with velocity $\dot{\varepsilon}_i l/\varepsilon_y$ cm sec^{-1} along the axis of the cylindrical specimen and, assuming the radial temperature differential is negligible, the heat flow equation becomes

$$\frac{\partial\theta}{\partial t}=\alpha\left(\frac{\mathrm{d}^2\theta}{\mathrm{d}x^2}-\frac{2h}{Kr}\theta\right), \qquad \ldots\ldots\ldots\ldots\ldots \quad \mathbf{(25)}$$

where θ is the temperature difference from the bath temperature T at position x at time t, α is the thermal diffusivity, K the thermal conductivity of steel, h a surface coefficient of heat transfer to the bath, and r is the radius of the specimen in appropriate units.

If, following Rosenthal (1946), a transformation is made from the fixed coordinate x to a moving coordinate x', where

$$x'=x-(\dot{\varepsilon}_i l/\varepsilon_y)t, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad \mathbf{(26)}$$

centred on the heat source and a quasi-stationary state is assumed such that $\partial\theta/\partial t$ is zero, then

$$-\frac{\dot{\varepsilon}_i l}{\alpha\varepsilon_y}\,\frac{\mathrm{d}\theta}{\mathrm{d}x'}=\frac{\mathrm{d}^2\theta}{\mathrm{d}x^2}-\frac{2h}{Kr}\theta, \qquad \ldots\ldots\ldots\ldots\ldots \quad \mathbf{(27)}$$

$$\left.\begin{array}{lll} \text{For } x'<0, & \theta=\theta_0\exp\{x'[-\dot{\varepsilon}_i l/2\alpha\varepsilon_y+\sqrt{\{(\dot{\varepsilon}_i l/2\alpha\varepsilon_y)^2+2h/Kr\}}]\}, \\ \phantom{\text{For }} x'>0, & \theta=\theta_0\exp\{-x'[\dot{\varepsilon}_i l/2\alpha\varepsilon_y+\sqrt{\{(\dot{\varepsilon}_i l/2\alpha\varepsilon_y)^2+2h/Kr\}}]\}, \\ \phantom{\text{For }} x'=0, & \theta_0=J\sigma_l\dot{\varepsilon}_i l/2K\sqrt{\{(\dot{\varepsilon}_i l/2\alpha\varepsilon_y)^2+2h/Kr\}}. \end{array}\right\} \qquad \ldots\ldots\ldots\ldots\ldots \quad \mathbf{(28)}$$

It will be noted as $\dot{\varepsilon}_i$ increases, for $-1\leqslant x<0$, when

$$\dot{\varepsilon}_i l/2\alpha\varepsilon_y>\sqrt{(2h/Kr)}, \qquad \ldots\ldots\ldots\ldots\ldots \quad \mathbf{(29)}$$

the temperature in the deformed region will approach θ_0.

It has been found in welding studies that a quasi-stationary temperature profile is obtained after the heat source has moved several centimetres at the normal welding speeds of 0·1–1 cm sec^{-1} (depending on the section) (Bornefeld 1933). At low temperatures, from the data of Kenyon (1948), α is greater by a factor of 10, and in our experiments in the soft machine the profile must have been established within distances of 10^{-1} cm, i.e. the velocity $\dot{\varepsilon}_i l/\varepsilon_y$ of the source would need to be of the order 10^2–10^3 cm sec^{-1} and (since at low temperatures the yield strain approaches 0·1, curve A, Fig. 1) the impressed strain rate of order 10–10^2 sec^{-1} to fully achieve a quasi-stationary temperature profile.

Appreciable temperature increase would, however, be expected at the strain rates of 1 to 10 sec^{-1} occurring during the yield drop. In the hard machine no effects due to temperature increase in the test piece would be expected to occur, since the deformation during the yield drop in load is smaller and the strain rate is lower than in the soft machine.

In the soft tensile tests carried out at the highest strain rate, $3\cdot 8\times 10^{-1}$ sec^{-1}, in *iso*pentane and propane with a moderate turbulence introduced by bubbling cold, dry air through the bath, from functions quoted by Jakob (1949) the Nusselt number would have been of order 10^2, i.e. h would have been nearly constant and independent of the testing temperature. In these tests the observed difference in yield stress approached 17×10^8 dynes cm^{-2} at low temperatures and the maximum strain rate during the yield load drop is estimated to be of order 10 sec^{-1}.

These values are sufficient to introduce adiabatic heating instabilities. It is significant that, as shown in Figure 4, although the ultimate stress decreased to the lower yield stress by −140 °C (a consequence of mechanical instability on a curve displaying moderate strain hardening), a maximum on the apparent lower yield stress curve was only observed at −171 °C when $\mathrm{d}\theta/\mathrm{d}T$, the theoretical rate of maximum adiabatic temperature increase, θ, in the specimen with temperature T approached −1, and stress strain curves of type E, Figure 1, were obtained. The apparent lower yield stress, in this case only, was defined as the stress at which the load drop was arrested. These specimens broke with a ductile fracture characteristic of failure at higher temperatures, whereas the specimens tested at −183 °C broke in a brittle manner at the upper yield stress before deformation-induced adiabatic heating could intervene.

In low temperature boiling liquids the surface coefficient of heat transfer h is not independent of the temperature and, from Farber and Scorah (1948), will vary as θ^y where y is also a function of θ increasing from 2 to 4 as θ increases from 0·1 to 50 °C.

Since the mechanism of heat transfer is basically the same, the variation of h for water is assumed to apply to liquefied gases and, in the temperature range of interest,

$$h = C\theta^2, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (30)$$

where C is 0·002 approximately. This relationship was checked by quenching cylinders of steel of two different diameters from liquid oxygen, and mixtures of liquid oxygen and nitrogen, into nitrogen and measuring the difference of the central temperature of the cylinders from the bath temperature. The variation of the derived mean value of h was consistent with the approximation.

In boiling liquids heat loss at the moving planar heat source will be very great, but away from the source heat losses from the surface of the cylinder will be lower than in the case where h is constant. It can then be assumed that temperatures deduced from inequality (**29**) and from (**30**) will be maintained over a considerable length of the deformed specimen

$$\theta<\frac{\dot{\varepsilon}_i l}{2\alpha\varepsilon_y}\sqrt{\left(\frac{Kr}{2C}\right)}, \qquad \textbf{(31)}$$

substituting $\alpha=1$, $K=0\cdot2$, $r=0\cdot125$, $l=2$, $\varepsilon_y=0\cdot1$,

$$\theta<12\dot{\varepsilon}_i. \qquad \textbf{(32)}$$

This agrees qualitatively with our estimate of the temperature increase required at strain rates of the order of 1 $\sec^{-1}$ to produce the serrations and does not predict any considerable temperature instability below $0\cdot1$ $\sec^{-1}$ strain rate. At higher strain rates the temperature will tend to be limited by the increase in exponent y of θ and by the temperature variation of the specific heat.

V. CONCLUSIONS

During the yield point load drop in steel the strain rate in the deforming region of the specimen increases rapidly. If the stress-strain curve displays little strain hardening and the inertial mass between the specimen and the extensible load recording device is too large, the displacement of the mass can produce a mechanical vibration at the lower yield point.

In normal tensile testing the difference between the upper and lower yield stresses is low and the effect is not observed. From (**21**) the effect will tend to be produced :

(i) If the applied strain rate is high, and a soft tensile testing machine is employed.

(ii) As the testing temperature is lowered, when the difference in yield stresses tends to increase and strain hardening after the yield strain to decrease.

(iii) If by careful alignment of the specimen, or by an aging treatment, the upper yield stress is increased while the lower yield stress is not significantly affected.

(iv) By the intervention of strain rate dependent phenomena, for example, crystallographic changes in the specimen or adiabatic temperature increase in the deformed zone, such that the effective lower yield stress is decreased during the load drop. The instability would render attempts to relate crystallographic changes to the resultant load drop somewhat difficult to interpret.

(v) If the ratio of length to area of the specimen is increased.

(vi) In face-centred cubic metals during the serrations observed at low temperatures and in the Portevin-Le Chatelier region, since these materials display little change in flow stress with increase in strain rate. On the other hand, the face-centred cubic metals work harden strongly and this would act to prevent the instability appearing.

In the particular case of yield instability investigated it appears improbable that the instability is of purely mechanical origin, although the observed relationship between the extent of the serration and the difference in the yield stresses, at two strain rates agrees, in form, with this hypothesis. It is suggested that the instability is aided by deformation-induced adiabatic heating which will not become important until the strain rate is great enough to minimize heat losses by conduction along the specimen and to the bath during the test. Corresponding to our observations, a critical difference in yield stresses is needed before the required strain rate can be attained in the deforming specimen during the yield point drop in load.

VI. Acknowledgments

The investigations on which this paper is based were carried out at l'Institut de Recherches de la Sidérurgie, St.-Germain-en-Laye, France. The author expresses his appreciation to IRSID for the provision of its unique laboratory facilities, and desires to thank Mr. C. Crussard and Mr. J. Plateau, who directed the research, for their continued interest and encouragement.

Mr. R. Eather carried out the numerical calculations from which Figures 8 and 9 were prepared. Permission by the Broken Hill Pty. Co. Ltd. to publish this paper is acknowledged.

VII. References

Allen, N. P., Hopkins, B. E., and McLennan, J. E. (1956).—*Proc. Roy. Soc.* A **234** : 221–46.

Basinski, Z. S. (1957).—*Proc. Roy. Soc.* A **240** : 229–42.

Blewitt, T. H., Coltman, R. R., and Redman, J. K. (1957).—*J. Appl. Phys.* **28** : 651–60.

Bornefeld, H. (1943).—*Techn. Zbl. practische Metallbearbeitung* **43** : 14.

Farber, E. A., and Scorah, R. L. (1948).—*Trans. Amer. Soc. Mech. Engrs.* **70** : 369, 652–6.

Fisher, J. C., and Rogers, H. C. (1956).—*Acta Met.* **4** : 180–5.

Jakob, M. (1949).—" Heat Transfer." Vol. 1. p. 563. (Chapman and Hall : London.)

Kenyon, R. L. (1948).—" Metals Handbook." (Ed. T. Lyman.) pp. 423–34. (American Society for Metals : Chicago.)

Kochendörfer, A., and Wink, W. (1957).—*Arch. Eisenhüttenw.* **28** : E 67–79.

Lean, J. B., and Plateau, J. (1958).—*C.R. Acad. Sci., Paris* **247** : 1196–9.

Lean, J. B., Plateau, J., and Crussard, C. (1958).—*C.R. Acad. Sci., Paris* **247** : 1458–61.

Lean, J. B., and Plateau, J. (1959).—*Mem. Sci. Rev. Met.* **56** (1) : 91–9.

McLachlan, N. W. (1956).—" Ordinary Non-linear Differential Equations." 2nd Edn. p. 145. (Oxford Univ. Press.)

Rosenthal, D. (1946).—*Trans. Amer. Soc. Mech. Engrs.* **68** : 849–65.

Uzhik, G. V. (1955).—*Izv. Akad. Nauk SSSR, Otdelenia Technicheskika Nauk* **1** : 57.

SMALL ANGLE INTENSITIES FROM DOUBLE BRAGG REFLECTIONS IN COLD-WORKED METALS*

By B. E. Warren†

[*Manuscript received September* 25, 1959]

Summary

The small angle intensity from a cold-worked polygrained metal is treated in terms of double Bragg reflections. It is assumed that cold work breaks up the original grains into subgrains with varying orientations, such that a first reflection in one subgrain can be followed by a second reflection in another subgrain of the original grain. For an initial sample which is polygrained with random orientation, the small angle intensity resulting from cold work is given in absolute units in terms of two parameters; an average initial grain size G, and a subgrain correlation function $p(\varphi)$ giving the distribution in orientation of subgrains with respect to any average subgrain of the original grain. Under favourable conditions, both of these quantities can be obtained from the small angle intensity curve. For polygrained metals, small angle scattering is a useful tool for obtaining the subgrain correlation function $p(\varphi)$, a quantity not obtainable from the usual high angle measurements.

I. INTRODUCTION

A small angle scattering from cold-worked metals has been observed by Blin and Guinier (1953). It was suggested by Beeman (Neynaber, Brammer, and Beeman 1959) that the small angle intensities could be due to double Bragg reflections. For any powder sample containing small grains with random orientation, there will be a small angle intensity resulting from two kinds of double Bragg reflections. Intergrain scattering consists of an *hkl* reflection from one grain, followed by the same *hkl* reflection in a second suitably oriented grain. Intragrain scattering involves an *hkl* reflection followed by a $\bar{h}\bar{k}\bar{l}$ reflection in the same small grain. For powder samples comprising grains with random orientation, the intergrain contribution is probably the more important. Expressed in electron units per atom, it is given by (Warren 1959)

$$I_{eua}=\frac{1}{2\theta^\circ}\left(\frac{e^4}{m^2c^4}\right)\frac{45\lambda^6v_0}{32\pi^3v_a^4\mu}\sum_i\frac{m_iF_i^4(1+\cos^4 2\theta_i)}{\sin^4\theta_i\sin^2 2\theta_i}\left\{1-\frac{(1-\exp[-g_i])}{g_i}\right\}, \quad \ldots \quad (1)$$

where

$g_i=2\mu t\sin^2\theta_i/|\cos 2\theta_i|$,
$2\theta^\circ$ is the scattering angle in degrees,
v_0 is the volume per atom,
v_a is the volume of the unit cell,
μ is the linear absorption coefficient,
Σ_i is a sum over all Bragg reflections up to $\sin\theta_i=1$,
m_i is the multiplicity,
F_i is the structure factor per unit cell,
t is the sample thickness.

---

* Research sponsored by the U.S. Atomic Energy Commission.

† Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.

The intergrain contribution for randomly oriented grains falls off inversely as $2\theta°$. Absolute measurements of the small angle intensity from cold-worked metals are considerably larger than the values predicted by equation (**1**), and they fall off more rapidly with increasing angle. Hence the contribution expressed by equation (**1**) is not the important effect. For cold-worked metals it appears that the important effect arises from a production of subgrains, and we now develop the appropriate theory.

II. Small Angle Intensities from Subgrain Clusters

We picture the undeformed sample as comprising grains with random orientation. The effect of cold work is to break up each of these original grains into a large number of subgrains having slightly varying orientations with respect to one another. If the primary beam makes a Bragg reflection in one of these subgrains, there is a high probability that the reflected beam will make a second Bragg reflection in the other subgrains of the original grain. This problem has been treated by Ogier, Wild, and Nickel (1959), and the theory which we shall develop here is simply an extension of their treatment.

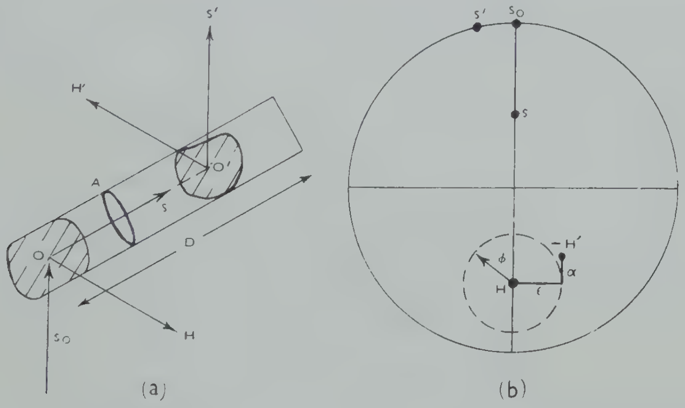


Fig. 1.—Relation between the first and second reflections in two subgrains which produce the small angle intensity from a cold-worked metal.

Figure 1 (*a*) shows the primary beam with direction $\vec{s}_0$ making a Bragg reflection on a subgrain at O. The reflected beam has a direction $\vec{s}$, intensity I, and cross sectional area A. The reflected beam makes a second Bragg reflection at O' on another subgrain of the original grain. The orientations of the three beams $\vec{s}_0$, $\vec{s}$, $\vec{s}'$ and of the planar normals $\vec{H}$ and $-\vec{H}'$ are represented by Figure 1 (*b*), which is a view of Figure 1 (*a*) looking from the right. The orientation of $-\vec{H}'$ relative to $\vec{H}$ is expressed by the two small angles ε, α perpendicular to one another.

Let $N(\varepsilon,\alpha)d\varepsilon d\alpha$ be the number of subgrains per unit volume in the original grain, with orientation between ε, $\varepsilon+d\varepsilon$ and α, $\alpha+d\alpha$ relative to any first reflecting subgrain such as that at O. Let D be the average length traversed by the first reflected beam through subgrains of the original grain, so that AD is the volume in the original grain irradiated by the first reflected beam. The number of subgrains with orientations between ε, $\varepsilon+\mathrm{d}\varepsilon$ and α, $\alpha+\mathrm{d}\alpha$ from the first reflecting subgrain, which are illuminated by the first reflected beam is then $ADN(\varepsilon,\alpha)\mathrm{d}\varepsilon\mathrm{d}\alpha$. If R is the sample to receiver distance, the intensity I' of the second reflected beam from one of the subgrains in the volume AD is given by

$$I'=I\frac{e^4}{m^2c^4R^2}F^2\frac{\sin^2(\pi/\lambda)(\vec{s}'-\vec{s})\cdot N_1\vec{a}_1}{\sin^2(\pi/\lambda)(\vec{s}'-\vec{s})\cdot\vec{a}_1}\,\frac{\sin^2(\pi/\lambda)(\vec{s}'-\vec{s})\cdot N_2\vec{a}_2}{\sin^2(\pi/\lambda)(\vec{s}'-\vec{s})\cdot\vec{a}_2}\,\frac{\sin^2(\pi/\lambda)(\vec{s}'-\vec{s})\cdot N_3\vec{a}_3}{\sin^2(\pi/\lambda)(\vec{s}'-\vec{s})\cdot\vec{a}_3}. \quad \ldots\ldots\ldots\ldots\ldots\ldots \ (2)$$

Let $\mathrm{d}P_1'$ be the diffracted power from all the grains in volume AD in the range ε, $\varepsilon+\mathrm{d}\varepsilon$ due to illumination by the first reflected beam I from subgrain 1

$$\mathrm{d}P_1'=ADN(\varepsilon,0)\mathrm{d}\varepsilon\int\!\!\int\!\!\int I'R^2\mathrm{d}\beta\mathrm{d}\gamma\mathrm{d}\alpha, \quad \ldots\ldots\ldots\ldots \ (3)$$

where $R^2\mathrm{d}\beta\mathrm{d}\gamma$ represents an area on the receiving surface, and $N(\varepsilon,\alpha)$ is replaced by $N(\varepsilon,0)$ since there is a contribution only for very small values of α. The integrals involved in equation (**3**) are those of the usual derivation of an integrated intensity, and hence we obtain

$$\mathrm{d}P_1'=DN(\varepsilon,0)\mathrm{d}\varepsilon(AI)\frac{e^4}{m^2c^4}\,\frac{F^2N^3\lambda^3}{v_a\sin 2\theta_i}, \quad \ldots\ldots\ldots\ldots \ (4)$$

where (AI) is the diffracted power from subgrain 1, N^3 is the average number of cells per subgrain, v_a is the cell volume, and θ_i is the Bragg angle.

For a fixed *hkl*, the total diffracted power in the second reflection due to all of the first reflections is obtained by summing over all first reflections:

$$\left.\begin{aligned}&\mathrm{d}P'=\Sigma\mathrm{d}P_1',\\&\Sigma(AI)=P=I_0\left(\frac{e^4}{m^2c^4}\right)\frac{\lambda^3mF^2A_0t}{4v_a^2\sin\theta_i},\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots \ (5)$$

where P is the total diffracted power in the first reflection for the whole sample, m is the *hkl* multiplicity, A_0 is the cross-sectional area of the primary beam so that I_0A_0 is the power P_0 in the primary beam. Let $\delta V=N^3v_a$ be the average volume per subgrain. Combining equations (**4**) and (**5**), introducing the proper polarization factor for a double reflection, and introducing an absorption factor which allows approximately for the extra path length in a double reflection, we obtain for $\mathrm{d}P'$

$$\mathrm{d}P'=P_0\left(\frac{e^4}{m^2c^4}\right)^2\frac{\lambda^6mF^4\delta Vt}{4v_a^4\sin\theta_i\sin 2\theta_i}\left(\frac{1+\cos^4 2\theta_i}{2}\right)\exp\left[-\mu(t+D\sin^2\theta_i)\right]DN(\varepsilon,0)\mathrm{d}\varepsilon. \quad \ldots\ldots\ldots\ldots\ldots\ldots \ (6)$$

The quantity $\mathrm{d}P'$ is the total diffracted power in a second *hkl* reflection due to all the subgrains within an original grain which are within ε, $\varepsilon+\mathrm{d}\varepsilon$ of the subgrain giving the first reflection. This power can be expressed in terms of a small angle intensity between angles 2θ and $2\theta+\mathrm{d}(2\theta)$

$$\mathrm{d}P'=I_{2\theta}2\pi R^2\sin 2\theta\mathrm{d}(2\theta). \quad \ldots\ldots\ldots\ldots\ldots \ (7)$$

From the relations shown by Figure 1, we can write $\vec{s}-\vec{s}_0=\lambda\vec{H}$, $\vec{s}'-\vec{s}=\lambda\vec{H}'$, and hence $|\vec{s}'-\vec{s}_0|=\lambda|\vec{H}-(-\vec{H}')|=\lambda|\vec{H}|\varepsilon$, from which

$$2\theta=(2\sin\theta_i)\varepsilon, \quad d(2\theta)=(2\sin\theta_i)d\varepsilon. \qquad \text{(8)}$$

Combining equations (**6**), (**7**), and (**8**), adding a factor of 2 to allow for contributions from $+\varepsilon$ and $-\varepsilon$, and summing over all *hkl* reflections up to $\sin\theta_i=1$, we obtain the total small angle intensity

$$I_{2\theta}=P_0\left(\frac{e^4}{m^2c^4}\right)^2 \frac{\lambda^6\delta V t\exp[-\mu t]D}{16\pi R^2 v_a^4\sin 2\theta}\sum_i\frac{m_iF_i^4(1+\cos^4 2\theta_i)N(\varepsilon,0)\exp[-\mu D\sin^2\theta_i]}{\sin^2\theta_i\sin 2\theta_i}. \qquad \text{(9)}$$

It is convenient to express both the theoretical and the experimental small angle intensities in absolute units, electron units per atom. Calling v_0 the volume per atom, an intensity in electron units per atom I_{eua} is defined by the equation

$$I_{2\theta}=I_0(e^4/m^2c^4R^2)I_{eua}(A_0t/v_0)\exp[-\mu t]. \qquad \text{(10)}$$

Combining equations (**9**), (**10**), we express the small angle intensity in absolute units

$$I_{eua}=\left(\frac{e^4}{m^2c^4}\right)\frac{\lambda^6\delta V D v_0}{16\pi v_a^4\sin 2\theta}\sum_i\frac{m_iF_i^4(1+\cos^4 2\theta_i)N(\varepsilon,0)\exp[-\mu D\sin^2\theta_i]}{\sin^2\theta_i\sin 2\theta_i}. \qquad \text{(11)}$$

Equation (**11**) gives the small angle intensity for a particular 2θ value, provided that for each *hkl* we use proper values of $N(\varepsilon,0)$ corresponding to the chosen 2θ value. From equation (**8**), $2\theta=(2\sin\theta_i)\varepsilon$ and hence for each *hkl* there is a different ε value for the chosen 2θ value. However, in the summation of equation (**11**), one reflection usually contributes more than all of the others combined, for example, the 111 reflection for face-centred cubic metals. Hence we can use the value of $N(\varepsilon,0)\exp[-\mu D\sin^2\theta_i]$ for the strongest contributor and take it outside of the summation.

We will assume that for purposes of a powder pattern reflection, the probability of a subgrain being rotated through an angle φ in any direction is the same as that of the rotation $(\varepsilon,0)$ where $\varphi=\varepsilon$. If $d\Omega$ is a small solid angle, we say that $N(\varphi)d\Omega=N(\varepsilon,0)d\Omega$. We now introduce a subgrain correlation function $p(\varphi)$

$$p(\varphi)=N(\varphi)\delta V=N(\varepsilon,0)\delta V, \qquad \text{(12)}$$

where $p(\varphi)d\Omega$ is the fraction of the subgrains in any original grain whose normals $-\vec{H}'$ fall in the solid angle $d\Omega$ at an angle φ to $\vec{H}$ of the average subgrain giving the first reflection. Expressing the scattering angle in degrees by $\sin 2\theta=(\pi/180)2\theta^\circ$, the small angle intensity is given by

$$I_{eua}(2\theta^\circ)=\frac{1}{2\theta^\circ}\,\frac{45}{4\pi^2}\left(\frac{e^4}{m^2c^4}\right)\frac{\lambda^6v_0[D'p(\varphi)]}{v_a^4}\sum_i\frac{m_iF_i^4(1+\cos^4 2\theta_i)}{\sin^2\theta_i\sin 2\theta_i}, \qquad \text{(13)}$$

where $D'=D\exp[-\mu D\sin^2\theta_j)$, $p(\varphi)$ refers to an angle φ such that $2\theta^\circ=(2\sin\theta_j)\varphi^\circ$, and the subscript j indicates the hkl reflection which contributes most to the sum.

For a sample of copper using Co $K\alpha$ radiation, the reflections 111, 200, 220, 311, 222, and 400 contribute to the sum in equation (**13**), but the contribution by 111 is greater than the others combined. Evaluated for the special case of a copper sample using Co $K\alpha$ radiation, equation (**13**) becomes

$$I_{eua}(2\theta^\circ)=\frac{49{,}200}{(2\theta^\circ)}[D'p(\varphi)], \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{14})$$

where D is in centimetres, $D'=D\exp[-133D]$ and $\varphi^\circ=1{\cdot}16(2\theta^\circ)$.

From equations (**13**) or (**14**), a set of measurements of $I_{eua}(2\theta^\circ)$ gives directly $[D'p(\varphi)]$ as a function of the angle φ°. To a good enough approximation, we can assume that D is about half of the original grain size G, and for the cases where D is sufficiently small we can write

$$D'=\tfrac{1}{2}G. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{15})$$

If values of the original grain size G are available from a previous metallographic examination, the measured values of $I_{eua}(2\theta^\circ)$ lead directly to the subgrain correlation function $p(\varphi)$ as a function of φ°.

If the values of $[D'p(\varphi)]$ obtained from $I_{eua}(2\theta^\circ)$ fall off sufficiently rapidly with φ°, we can make use of the normalization condition

$$\int p(\varphi)d\Omega=\int p(\varphi)2\pi\varphi^\circ(\pi/180)^2d\varphi^\circ=1,$$

and hence

$$D'=\frac{1}{540}\int[D'p(\varphi)]\varphi^\circ d\varphi^\circ. \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{16})$$

Hence, if $[D'p(\varphi)]\varphi^\circ$ falls off rapidly enough with φ° to determine the area under the curve, we can obtain both the effective original grain size $G=2D$, and absolute values of the subgrain correlation function $p(\varphi)$, directly from the small angle intensity curve.

III. Absolute Small Angle Intensities by Comparison with " Lucite "

It is important to have the small angle intensities in absolute units, and of the various methods for accomplishing this we shall discuss just one. The small angle intensities can be obtained in absolute units by comparison with the scattering from a sheet of commercial " Lucite ", $C_5H_8O_2$, of thickness 1·70 mm. Using Co $K\alpha$ radiation, a diffractometer trace from a thick sheet of " Lucite " shows a peak at $2\theta=16^\circ$ but no detail at the higher angles. With balanced Fe-Mn filters, the intensity ratio was determined $I(16^\circ)/I(160^\circ)=8{\cdot}35$. Designating the unit $C_5H_8O_2$ as a molecule, the intensity in electron units per molecule at 160° is given closely by the independent unmodified and modified scattering which can be obtained from the tables

$$I_{eum}(160^\circ)=\sum_m(f^2+Z-\Sigma f_e^2)=66.$$

The intensity in electron units per molecule at $2\theta = 16°$ is then given by

$$I_{eum}(16°) = 66 \times 8{\cdot}35(1+\cos^2 160)/(1+\cos^2 16) = 540.$$

With fixed slits, the counting rate is proportional to the intensity arriving at the receiving slit $CR = KI$, and the counting rate from " Lucite " at 16° is given by

$$(CR)_{\mathrm{L}16°} = \frac{KI_0e^4}{m^2c^4R^2}\left(\frac{1+\cos^2 16°}{2}\right) I_{eum}(16°)\frac{A_0t}{v_M} \exp\left[-\mu t'\right],$$

where A_0 is the cross-sectional area of the primary beam, t is the sample thickness, and v_M is the volume per molecule, so that A_0t/v_M is the effective number of molecules in the sample. For scattering at 16° the effective thickness is t'. With the values $t = 0{\cdot}170$ cm, $v_M = 141 \times 10^{-24}$ cm^3, and $\exp[-\mu t'] = 0{\cdot}157$, we obtain

$$(CR)_{\mathrm{L}16°} = 9{\cdot}8 \times 10^{22}(KA_0I_0e^4/m^2c^4R^2). \quad \ldots\ldots\ldots\ldots \quad \mathbf{(17)}$$

For the metal sample of thickness t and volume per atom v_0, the counting rate can be expressed in terms of I_{eua} the intensity in electron units per atom

$$(CR)_M = \frac{KI_0e^4}{m^2c^4R^2} I_{eua} \frac{A_0t}{v_0} \exp\left[-\mu t\right]. \quad \ldots\ldots\ldots\ldots \quad \mathbf{(18)}$$

Combining equations (**17**) and (**18**), we obtain

$$I_{eua} = 9{\cdot}8 \times 10^{22}\frac{v_0}{t\exp\left[-\mu t\right]} \cdot \frac{(CR)_M}{(CR)_{\mathrm{L}16°}}. \quad \ldots\ldots \quad \mathbf{(19)}$$

IV. Application to Copper Sheet

As an example, we can use a set of measurements of the small angle intensity from a sheet of commercial copper, cold worked by pounding down from 0·012 to 0·004 cm thickness. The measurements were made using Co $K\alpha$ radiation and Geiger counter recording with balanced Fe-Mn filters at the receiving slit. The measurements were converted to absolute units by comparison with the scattering from " Lucite ".

Figure 2 shows the measured small angle intensity for the copper sheet in the cold-worked state, and after annealing 2 hr at 400 °C. Curve C of Figure 2 shows the intergrain contribution for undeformed grains with completely random orientation computed from equation (**1**). It is evident that double Bragg reflections in undeformed grains with completely random orientations are quite inadequate to account for the small angle intensity from the cold-worked copper.

To allow roughly for extraneous scattering, we take the difference between the cold-worked and the annealed curves as representing the small angle intensity due to cold work. By use of equation (**14**), we then obtain the quantity $[D'p(\varphi)]$ shown by curve A of Figure 3. The quantity $[D'p(\varphi)]\varphi°$ is shown by curve B of Figure 3. From the area under this curve, and the use of equation (**16**), we obtain an approximate value $D' = 0{\cdot}35 \times 10^{-3}$ cm. An approximate size of the regions, within which the subgrains are part of an original grain, is then given by $G=2D'=0\cdot7\times10^{-3}$ cm. This value is not inconsistent with the sample thickness $t=4\cdot0\times10^{-3}$ cm. However, very little significance should be attached to these numbers, and they are given merely by way of illustrating a procedure. The initial state and texture of the sample is unknown, there was no metallographic examination of the original grain size, and the pounding down of the sample from 0·012 to 0·004 cm thickness was probably too drastic a cold work for the optimum utilization of the theory which has been developed.

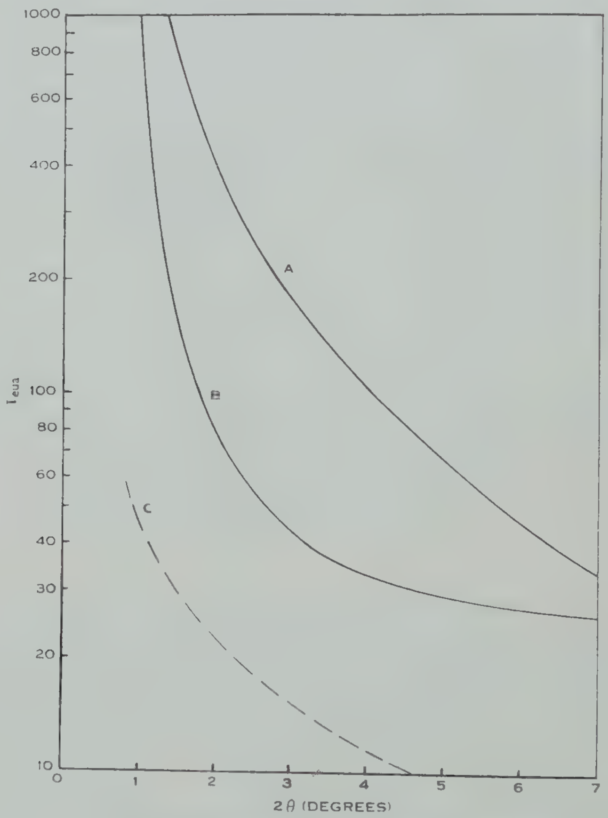


Fig. 2.—Small angle intensity in electron units per atom for a sheet of copper pounded down from 0·012 to 0·004 cm thickness (Co $K\alpha$). *A*, cold worked; *B*, after 2 hr anneal at 400 °C; *C*, computed double Bragg intensity for undeformed grains with random orientation.

V. CONCLUSIONS

The principal conclusion has already been given by Ogier, Wild, and Nickel (1959). Small angle intensity measurements from a cold-worked polygrained sample are capable of giving a subgrain correlation function $p(\varphi)$ representing the distribution in tilts of the subgrains relative to any average subgrain of the original grain. For cold-worked polygrained samples, this kind of information cannot be obtained from the usual large angle measurements, which give information about the individual subgrains, but not about their relative disorientation. The theory which has been developed here applies to a sample which, before the cold work, consisted of small grains with random orientation. Appropriate changes in the theory are of course required for samples in which there is a high degree of preferred orientation.

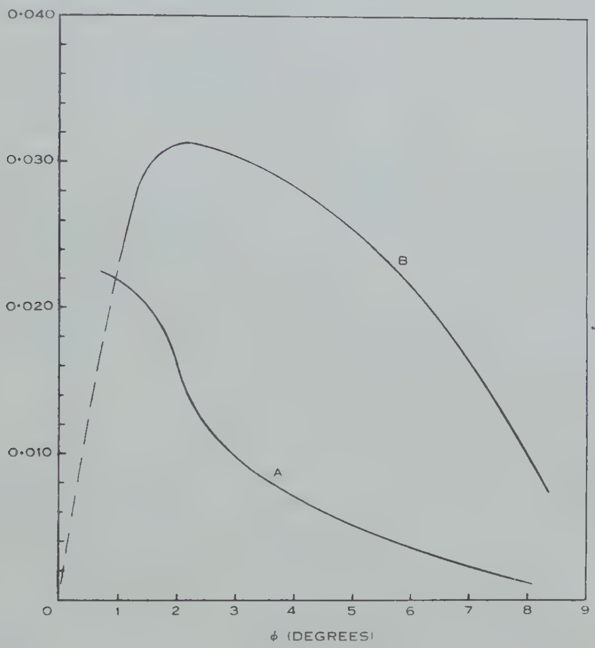


Fig. 3.—The subgrain correlation function $p(\varphi)$ for a sheet of copper pounded down from 0·012 to 0·004 cm thickness. Curve *A* gives $[D'p(\varphi)]$, curve *B* gives $[D'p(\varphi)]\varphi°$.

VI. Acknowledgments

I am indebted to Dr. R. L. Wild for making available in advance a copy of the paper by Ogier, Wild, and Nickel, which gave the first treatment of this problem, and for which the present paper is simply an extension.

VII. References

Blin, J., and Guinier, A. (1953).—*C.R. Acad. Sci., Paris* **236**: 2150.
Neynaber, R. H., Brammer, W. G., and Beeman, W. W. (1959).—*J. Appl. Phys.* **30**: 656.
Ogier, W. T., Wild, R. L., and Nickel, J. C. (1959).—*J. Appl. Phys.* **30**: 408.
Warren, B. E. (1959).—*Acta Cryst.* **12**: 837.

X-RAY MEASUREMENT OF TWIN FAULTING IN FACE-CENTRED CUBIC METALS*

By B. E. Warren†

[*Manuscript received September 23, 1959*]

Summary

It was shown by Paterson that in a cold-worked face-centred cubic metal, deformation faults produce a peak shift, and twin faults produce a peak asymmetry. If the peak shape is represented by a Fourier series, the asymmetry is expressed by the sine terms. However, the evaluation of the sine coefficients is difficult, due to overlapping of the tails of neighbouring reflections and uncertainty in the peak origin. In this paper, a method is developed which combines the tails of the (111) and (200) reflections. This allows a determination of the twin fault probability in face-centred cubic metals which is more nearly independent of overlapping of the tails and choice of the peak origin.

I. Introduction

The possibility of determining deformation fault and twin fault probabilities from the powder patterns of cold-worked face-centred cubic metals was demonstrated in the pioneer paper by M. S. Paterson (1952). It was shown that in addition to a peak broadening, deformation faults produced a peak shift and twin faults produced a peak asymmetry. It is a relatively easy matter to determine the deformation fault probability from the measured peak shifts, but the determination of the twin fault probability from the peak asymmetry is not as simple.

If the peak shape is expressed by a Fourier series, the asymmetry is represented by the sine terms, and the twin fault probability can be obtained from the sine coefficients, as was done by Wagner (1957*a*, 1957*b*) for several cold-worked face-centred cubic metals. However, it is not a simple matter to determine the sine coefficients, since the asymmetry is usually small and in addition there is overlapping of the tails of neighbouring reflections and an uncertainty in the position of the peak origin. It is the purpose of this paper to develop a method for determining the twin fault probability in face-centred cubic metals, which is more nearly independent of the overlapping of tails and the uncertainty in the position of the peak origin.

II. Theory of the Peak Asymmetry

We consider the powder sample of a cold-worked face-centred cubic metal in which α and β are the probabilities of deformation faults and twin faults on the 111 planes. We consider α and β to be independent of one another and both small. It is convenient to introduce hexagonal axes $\vec{A}_1\,\vec{A}_2\,\vec{A}_3$, where $\vec{A}_1$ and $\vec{A}_2$

---

* Research sponsored by the U.S. Atomic Energy Commission.

† Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.

are in the 111 plane, and $\vec{A}_3$ is normal to the 111 plane. The hexagonal axes and indices are given by the transformations:

$$\left.\begin{array}{ll} \vec{A}_1=-\vec{a}_1/2+\vec{a}_2/2+0, & H=-h/2+k/2+0, \\ \vec{A}_2=0-\vec{a}_2/2+\vec{a}_3/2, & K=0-k/2+l/2, \\ \vec{A}_3=\vec{a}_1+\vec{a}_2+\vec{a}_3, & L_0=h+k+l. \end{array}\right\} \quad \ldots\ldots (1)$$

The diffraction vector is represented in terms of the continuous variables $h_1h_2h_3$ and the vectors $\vec{B}_1\vec{B}_2\vec{B}_3$ reciprocal to the $\vec{A}_1\vec{A}_2\vec{A}_3$,

$$(\vec{s}-\vec{s}_0)/\lambda=h_1\vec{B}_1+h_2\vec{B}_2+h_3\vec{B}_3. \quad \ldots\ldots\ldots\ldots (2)$$

If N_1 and N_2 are the number of repetitions in the $\vec{A}_1$ and $\vec{A}_2$ directions, the intensity from a single 111 layer is given by

$$\psi^2=I_ef^2\frac{\sin^2(\pi/\lambda)(\vec{s}-\vec{s}_0)\cdot N_1\vec{A}_1}{\sin^2(\pi/\lambda)(\vec{s}-\vec{s}_0)\cdot\vec{A}_1}\,\frac{\sin^2(\pi/\lambda)(\vec{s}-\vec{s}_0)\cdot N_2\vec{A}_2}{\sin^2(\pi/\lambda)(\vec{s}-\vec{s}_0)\cdot\vec{A}_2}. \quad \ldots (3)$$

For $L_0=3N$ there is no effect by faulting. For the case $L_0=3N\pm1$ where faulting plays a role, we can combine Paterson's results in which deformation and twin faulting were treated separately. To do this, we use a simplification of a combined treatment given by Gevers (1954). Omitting the preliminary steps which have been given elsewhere (Warren 1959), the interference function is given by

$$I(h_1h_2h_3)=\psi^2\sum_{m=-\infty}^{+\infty}N_mZ^{|m|}\left\{\cos 2\pi m\left(\frac{h_3-L_0}{3}-(\pm)\frac{3^{\frac{1}{2}}\alpha}{4\pi}\right)\right.$$
$$\left.\pm\frac{\beta}{3^{\frac{1}{2}}}\sin 2\pi\,|\,m\,|\left(\frac{h_3-L_0}{3}(\pm)\frac{3^{\frac{1}{2}}\alpha}{4\pi}\right)\right\}, \quad \ldots\ldots\ldots (4)$$

where N_m is the number of 111 layers having an mth neighbour, $Z=1-1{\cdot}5\alpha-\beta$, and the choice in sign $(\pm)$ corresponds to $L_0=3N\pm1$.

The observable quantity is $P'_{2\theta}$ where $P'_{2\theta}\mathrm{d}(2\theta)$ is the diffracted power between 2θ and $2\theta+\mathrm{d}(2\theta)$ per unit length of diffraction circle. We now convert to $P'_{2\theta}$ using the powder pattern power theorem, and the connecting relation $\mathrm{d}(2\theta)=\lambda\,|\,\vec{B}_3\,|\sin\varphi\,\mathrm{d}h_3/\cos\theta$, where $\sin\varphi=L_0\,|\,\vec{B}_3\,|\,d$, and d is the planar spacing. Introduce the abbreviation

$$G=I_eN_0R\lambda^2f^2/16\pi v_0\sin^2\theta, \quad \ldots\ldots\ldots\ldots\ldots (5)$$

where $N_0=MN_1N_2N_3$ is the number of atoms in the sample, N_3 is the average number of 111 layers per crystal, and v_0 is the volume per atom. For the single reflection HKL_0 where $L_0=3N\pm1$, the distribution with respect to 2θ is given by

$$P'_{2\theta}(HKL_0)=\frac{G}{3|\,\vec{B}_3\,||\sin\varphi\,|}\sum_{m=-\infty}^{+\infty}\frac{N_mZ^{|m|}}{N_3}\left\{\cos 2\pi m\left[\frac{h_3-L_0}{3}-(\pm)\frac{3^{\frac{1}{2}}\alpha}{4\pi}\right]\right.$$
$$\left.\pm\frac{\beta}{3^{\frac{1}{2}}}\sin 2\pi|\,m\,|\left[\frac{h_3-L_0}{3}-(\pm)\frac{3^{\frac{1}{2}}\alpha}{4\pi}\right]\right\}. \quad \ldots\ldots\ldots\ldots (6)$$

To analyse the peak shape in terms of Fourier coefficients, it is convenient to consider each reflection as $00l'$ in terms of orthorhombic axes $\vec{a}'_1\vec{a}'_2\vec{a}'_3$ and corresponding reciprocal axes $\vec{b}'_1\vec{b}'_2\vec{b}'_3$. The connecting relations are $(2\sin\theta)/\lambda = h'_3|\vec{b}'_3|$, $(h_3-L_0)|\vec{B}_3|\sin\varphi = (h'_3-l')|\vec{b}'_3|$.

With the abbreviations

$$q=|\vec{b}'_3|/[3|\vec{B}_3|\sin\varphi], \qquad \varepsilon=3^{\frac{1}{2}}\alpha/4\pi q, \qquad \ldots\ldots\ldots \quad (7)$$

Equation (6) becomes

$$P'_{2\theta}(HKL_0)=\frac{G}{|\vec{b}'_3|}\sum_{n=-\infty}^{+\infty}\frac{N(|n/q|)}{N_3}Z^{|n/q|}\Big\{\cos 2\pi n[h'_3-l'-(\pm)\varepsilon] \pm\frac{\beta}{3^{\frac{1}{2}}}\frac{L_0}{|L_0|}\sin 2\pi|n|[h'_3-l'-(\pm)\varepsilon]\Big\}, \qquad \ldots\ldots\ldots\ldots \quad (8)$$

where the summation is over all integral values of the new index n.

Each powder pattern reflection contains several HKL_0 components. Some of the components are broadened by faulting and the others are not. We will designate by b the components $L_0=3N\pm1$ which are broadened by faulting, and by u the components $L_0=3N$ which are unbroadened by faulting. The contribution by a broadened component is given by equation (8), and that of an unbroadened component by equation (8) with $Z=1$, $\varepsilon=0$, and $\beta=0$. The total contribution to a face-centred cubic powder pattern reflection $P'_{2\theta}(h_0)$, where $h_0^2=h^2+k^2+l^2$, is then given by a sum over the broadened and unbroadened components

$$P'_{2\theta}(h_0)=\frac{G}{|\vec{b}'_3|}\sum_{n=-\infty}^{+\infty}\Big\{\sum_b\frac{N(|n/q|)}{N_3}Z^{|n/q|}\cos 2\pi n[h'_3-l'-(\pm)\varepsilon] +\sum_u\frac{N(|n/q|)}{N_3}\cos 2\pi n(h'_3-l')+\frac{\beta}{3^{\frac{1}{2}}}\sum_b\frac{N(|n/q|)}{N_3}Z^{|n/q|}(\pm)\frac{L_0}{|L_0|}\sin 2\pi|n| [h'_3-l'-(\pm)\varepsilon]\Big\}. \qquad \ldots\ldots\ldots\ldots \quad (9)$$

We now make the approximation that the two cosine terms can be replaced by a single cosine term having an average coefficient and an average displacement. It is a good enough approximation to use the same displacement for the sine terms as for the cosine terms. Equation (9) is then expressed in the simple form

$$P'_{2\theta}(h_0)=\frac{G(u+b)}{|\vec{b}'_3|}\sum_{n=-\infty}^{+\infty}\{A_n\cos 2\pi n(h'_3-l'-\delta)+B_n\sin 2\pi|n|(h'_3-l'-\delta)\}, \qquad \ldots\ldots\ldots\ldots \quad (10)$$

where

$$\delta = [\sum_b (\pm)\varepsilon]/(u+b),$$

$$A_n = 1 - \frac{|n|}{|\vec{b}_3'|}\left[\frac{1}{D} + \frac{(1\cdot5\alpha+\beta)}{a(u+b)h_0} \sum_b |L_0|\right],$$

$$B_n = \frac{\beta}{3^{\frac{1}{2}}(u+b)} \sum_b (\pm)\frac{L_0}{|L_0|} \frac{N(|n/q|)}{N_3} Z^{|n/q|}.$$

The cosine coefficient A_n has been approximated for values of n which are small, $1/D = \langle \sin\varphi/N_3 d_{111}\rangle$ is an averaged coherent domain dimension in a direction normal to the reflecting planes, and a is the edge of the cubic cell.

The peak shift due to deformation faulting is expressed by the quantity δ. The particle size broadening which results from coherent domain size D and also from deformation and twin faulting is contained in the cosine coefficient A_n. If reliable values of the sine coefficients B_n were available, the twin fault probability β could be obtained from the limiting value of B_n as n approaches zero

$$(B_n)_{n\to 0} = \frac{\beta}{3^{\frac{1}{2}}(u+b)} \sum_b (\pm)\frac{L_0}{|L_0|}. \qquad \ldots\ldots\ldots\ldots (11)$$

However, we are interested in developing a method for determining β which is more nearly independent of the difficulties in determining B_n which arise from the choice in origin position and the overlapping of the tails of neighbouring reflections.

For the low order reflections 111, 200, 220, 222, and 400, the values of $|L_0|$ and of $(\pm)L_0/|L_0|$ are the same for all of the broadened components, and in the expression for B_n the sum over b can be replaced by multiplication. For small n we can approximate $N(|n/q|)$ and write

$$\frac{N(|n/q|)}{N_3} Z^{|n/q|} = \left(1 - \frac{|n|}{|\vec{b}_3'|D}\right)(1 - 1\cdot5\alpha - \beta)^{|nL_0/b_3' a h_0|}.$$

If the coherent domain size D is large enough so that the decrease in B_n due to D is small compared with that due to faulting, we can represent the sine coefficients for small n by

$$B_n = \frac{\beta b}{3^{\frac{1}{2}}(u+b)}(\pm)\frac{L_0}{|L_0|}Q^{|n|}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots (12)$$

where

$$Q = 1 - \left[\frac{1}{D|\vec{b}_3'|} + \frac{|L_0|}{|\vec{b}_3'|ah_0}(1\cdot5\alpha+\beta)\right].$$

The contribution to the observed peak, which arises from the sine terms, is then given by

$$P_{2\theta}'(s) = \frac{G\beta b}{3^{\frac{1}{2}}|\vec{b}_3'|}(\pm)\frac{L_0}{|L_0|} \sum_{n=-\infty}^{+\infty} Q^{|n|} \sin 2\pi|n|x, \qquad \ldots\ldots (13)$$

where $x=h_3'-l'-\delta$. Although the terms in equation (**13**) are exact only for small values of n, the expression will be used only at large values of x where the effect of weak high-order sine terms averages to zero. Evaluating the sum

$$P_{2\theta}'(s)=\frac{G\beta b}{3^{\frac{1}{2}}|\vec{b}_3'|}(\pm)\frac{L_0}{|L_0|}\;\frac{1}{c\tan\pi x}, \quad \ldots\ldots\ldots\ldots \textbf{(14)}$$

where

$$c=1+\left(\frac{1-Q}{2Q^{\frac{1}{2}}\sin\pi x}\right)^2$$

is a correction factor which can be evaluated if necessary, but which to a first approximation can be set equal to unity if the measurements are made at large enough values of x (for example $x=\frac{1}{4}$).

The factor G can be evaluated from the area of the peak. For this it is convenient to choose an interval within which the peak shows measurable variation similar to the interval chosen for evaluating Fourier coefficients. If the peak occurs at $2\theta_0$, the interval extends from $2\theta_1$ to $2\theta_2$ where $\sin\theta_2-\sin\theta_0=\sin\theta_0-\sin\theta_1$. The interval is divided into 60 subintervals, and positions x are represented by $x=i/60$ where i runs from -30 to $+30$. Since

$$(l'+\tfrac{1}{2})|\vec{b}_3'|=(2\sin\theta_2)/\lambda \text{ and } l'|\vec{b}_3'|=(2\sin\theta_0)/\lambda,$$

it follows that $|\vec{b}_3'|=4\cos\theta(\theta_2-\theta_0)/\lambda$. Combining this with the angular dependent part of G, we can write

$$G/|\vec{b}_3'|=G'/\{p(\theta_2-\theta_0)\},$$

where G' is a constant and p is an angular factor

$$p=\sin\theta\sin 2\theta/\{f^2(1+\cos^2 2\theta)\}. \quad \ldots\ldots\ldots \textbf{(15)}$$

Calling Y_i the peak ordinate at position i, in taking the difference of the corrected ordinates at positions equidistant on either side of the peak, the contribution by the cosine terms cancels out, and equation (**14**) can be put into the form

$$p_iY_i-p_{-i}Y_{-i}=\frac{2G'\beta b}{3^{\frac{1}{2}}(\theta_2-\theta_0)}(\pm)\frac{L_0}{|L_0|}\;\frac{1}{c_i\tan|\pi i/60|}. \quad \ldots \textbf{(16)}$$

The asymmetry effect is large for (111) and (200) and opposite in sign, so that both peaks have high tails between (111) and (200) and low tails outside of the peaks. Due to the overlapping of the tails, it is better to combine equations for (111) and (200). Let i stand for intervals in the (111) reflection, and j for intervals in the (200) reflection, and let the interval $\theta_2-\theta_0$ have the same magnitude for both reflections. The combination gives

$$(p_iY_i-p_{-i}Y_{-i})-(p_jY_j-p_{-j}Y_{-j})=\frac{2G'\beta}{3^{\frac{1}{2}}(\theta_2-\theta_0)}\left\{\left[\frac{(\pm)bL_0/|L_0|}{c_i\tan|\pi i/60|}\right]_{111}-\left[\frac{(\pm)BL_0/|L_0|}{c_j\tan|\pi j/60|}\right]_{200}\right\}.$$

It is convenient to evaluate the constant $G'/(\theta_2-\theta_0)$ from the area of the (200) peak since this is lower than (111),

$$\int_{-\frac{1}{2}}^{+\frac{1}{2}} P'_{2\theta}\mathrm{d}x=\frac{G'(u+b)}{p(\theta_2-\theta_0)}\int_{-\frac{1}{2}}^{+\frac{1}{2}} \sum_n A_n \cos 2\pi n x \mathrm{d}x=\frac{G'(u+b)}{p(\theta_2-\theta_0)},$$

and hence

$$\frac{1}{60}\sum_{j=-30}^{+30} Y_j=\frac{G'}{(\theta_2-\theta_0)}\left(\frac{u+b}{p}\right)_{200}.$$

To eliminate the trouble by overlapping between (111) and (200), we make position i coincide with position $-j$, so that it is only necessary to measure the total ordinate Y_i+Y_{-j} at this position. For (111) and (200) the values of $\sum_b(\pm)L_0/|L_0|$ are $+6$ and -6 respectively. Using these values, we obtain a final equation for determining β,

$$p_i(Y_i+Y_{-j})-(p_{-i}Y_{-i}+p_jY_j)=\frac{\beta p(200)}{30\times 3^{\frac{1}{2}}}\left\{\frac{1}{c_i \tan|\pi i/60|}+\frac{1}{c_j \tan|\pi j/60|}\right\}\sum_{j=-30}^{+30} Y_j. \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(17)}$$

The correction factors p are given by equation (**15**), and $p(200)$ is evaluated at the centre of the (200) peak. If the position i is taken approximately halfway between (111) and (200), it is a good first approximation to set c_i and c_j both equal to unity.

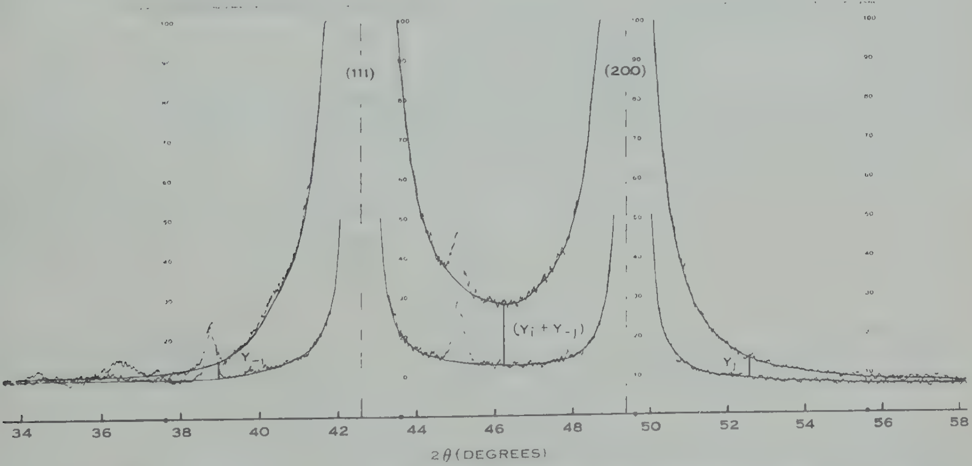


Fig. 1.—Recordings at room temperature of the (111) and (200) reflections from cold-worked and annealed filings of 70/30 α brass, using crystal monochromated Cu $K\alpha$ radiation. The filings were made under liquid nitrogen, and the annealing was 1 hr at 450 °C. Small peaks at 38·8° and 45° are from traces of LiF put on the sample surfaces for 2θ calibration purposes.

III. Application to 70/30 α Brass

To illustrate the application of equation (**17**), we use a sample of 70/30 α brass cold worked by making filings under liquid nitrogen. Figure 1 shows the diffraction pattern made at room temperature with crystal monochromated Cu $K\alpha$ radiation. To locate correctly the background level, a trace is also made under identical conditions using filings annealed 1 hr at 450 °C. Using crystal monochromated radiation, the background for the annealed specimen is due only to Compton modified and temperature diffuse scattering, and is hence independent of any extinction in the annealed sample. The position $(i, -j)$ is chosen approximately midway between the (111) and (200) peaks, and the corresponding positions $-i$ and $+j$ equidistant from the centres of the cold-worked peaks. From the $(\theta_2-\theta_0)$ interval which was chosen, $i=18\cdot5$ and $j=16$. The ordinates Y_{-i}, (Y_i+Y_{-j}), and Y_j are shown on the figure. A second tracing, with scale reduced by a known factor, is used to obtain the area of the (200) peak. This involves separating the overlapping of the tails between (111) and (200), but the peak area is not very sensitive to this separation. Using equation (**17**) and the values from Figure 1, we obtain $\beta=0\cdot07$. Since equation (**17**) involves a difference, it is doubtful if β can be determined to better than about $\pm0\cdot01$, and hence $0\cdot07$ is in satisfactory agreement with Wagner's (1957*a*) earlier value $\beta=0\cdot06$.

The method which has been outlined for determining twin fault probabilities in cold-worked face-centred cubic powder samples has several advantages: it does not involve any Fourier analysis, it allows automatically for overlapping of the tails, and it is relatively insensitive to choosing the centre of the cold work peaks. However, it is doubtful if an accuracy in β of better than $\pm0\cdot01$ can be expected.

IV. REFERENCES

GEVERS, R. (1954).—*Acta Cryst.* **7**: 337.
PATERSON, M. S. (1952).—*J. Appl. Phys.* **23**: 805.
WAGNER, C. N. J. (1957*a*).—*Acta Metall.* **5**: 427.
WAGNER, C. N. J. (1957*b*).—*Acta Metall.* **5**: 477.
WARREN, B. E. (1959).—X-ray studies of deformed metals. *Progr. Metal Phys.* **8**: 147.

FIELD EMISSION THROUGH DIELECTRIC LAYERS

By R. Gomer*

[*Manuscript received September 25, 1959*]

Summary

Field emission experiments on the adsorption of inert gases on tungsten have shown the presence of stepwise multilayer adsorption. It is possible to obtain emission data through these layers. It turns out that after suitable corrections for the reduction in field by dielectric effects a thinning of the field emission barrier remains. This is attributed to the fact that individual inert gas atoms represent short-range but deep potential holes (as in ordinary scattering) and so make the barrier more transparent. It is possible to treat the effect quantitatively in two approximations. One is due to Fermi and yields scattering lengths. The other consists of neglecting electronic motion transverse to the barrier, thereby making the problem one-dimensional and results in potential well dimensions. It is possible to calculate theoretical Fermi lengths for various assumed well shapes and dimensions. This is done for square wells and used to compare the results of the two approximations, in terms of the effective nuclear charges required to produce matching of the experimental Fermi lengths with the experimental well dimensions. Reasonable values of Z result for neon, argon, krypton, and xenon. Since the zero of the potential wells lies above that of the tunnelling electrons these experiments represent in fact scattering of negative energy electrons.

I. Introduction

The adsorption of inert gases on tungsten field emitters has been studied independently by Ehrlich (1959) and by the author (1958, 1959), with almost complete experimental agreement. This paper describes and attempts to interpret some of the phenomena encountered, with emphasis on those connected with emission through several molecular layers.

A brief description of the technique and the general results will first be given.

II. Method

The experiments consist of evaporating inert gases onto a tungsten field emitter, generally in such a way that only a portion of the latter receives a gas deposit ; subsequent heating makes it possible to observe diffusion and finally evaporation. During the experiment the emission tube is immersed in liquid helium or hydrogen. This provides the necessary high vacuum and makes unilateral gas deposition possible, since the sticking coefficient of atoms on the cold walls of the tube is very high. The gas source consists of a heatable platinum strip or tube on which the gas contained in the tube is precondensed at the start of the experiment (Gomer 1959).

Emission data were obtained by using stabilized power supplies, precision potentiometers for voltage determinations, and a vibrating reed electrometer

---

* Institute for the Study of Metals and Department of Chemistry, University of Chicago, Chicago, Illinois, U.S.A.

for current measurements. Details of the procedures have been given elsewhere (Gomer, Wortman, and Lundy 1957). Work functions were determined by using the Fowler-Nordheim equation (Good and Müller 1955)

$$I/V^2 = b \exp -6 \cdot 8 \times 10^7 \varphi^{3/2} \alpha / kV, \quad \ldots\ldots\ldots\ (1)$$

and comparing the slopes S of $\log (I/V^2)$ versus $1/V$ plots:

$$\varphi_G = (S_G/S_W)^{2/3} (\alpha_W/\alpha_G)^{2/3} \times 4 \cdot 5. \quad \ldots\ldots\ldots\ (2)$$

In these equations I is the total emitted current, V the voltage, b a constant, φ the work function, α the Nordheim image correction factor (Good and Müller 1955), and

$$k = F/V,$$

F being the electric field. The subscripts G and W refer to quantities pertaining to gas-covered or clean emitters respectively.

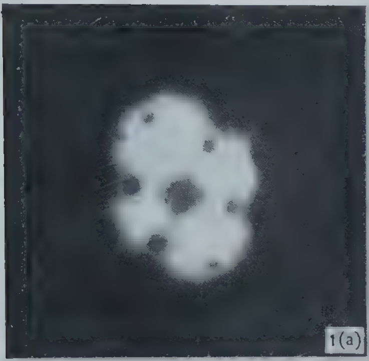
III. Summary of Adsorption Results

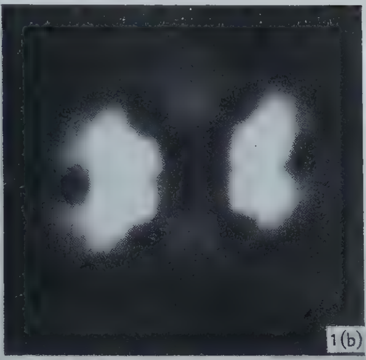
Deposits of a monolayer or less enhance emission, most markedly in the regions surrounding the 100 faces (Plate 1, Figs. 1 (*a*), 1 (*b*)). Diffusion occurs without boundaries under these conditions in the appropriate temperature ranges (Gomer 1958, 1959). When larger amounts of gas are evaporated onto the tip the region of initial deposit is darkened. Boundary free diffusion sets in at the same temperatures as before and again results in enhanced emission on the originally clean portions of the emitter. Further slight increases in temperature, however, lead to another type of diffusion in this case, consisting of waves of adsorbate flowing over the tip with sharply defined boundaries. It is possible to see several of these simultaneously inundating the surface (Plate 1, Fig. 2). Each succeeding wave or layer causes a decrease in emission and some loss in resolution. The appearance of a typical multilayer pattern is shown in Plate 1, Figure 3. It is seen that the emission anisotropy is unchanged from that of the monolayer case, although the pattern is considerably blurred.

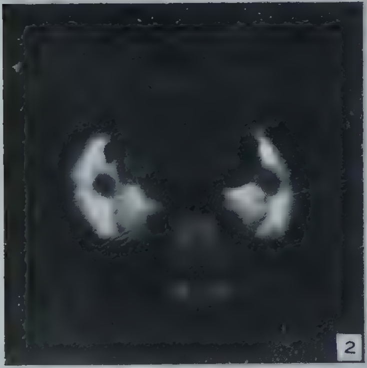
Further heating leads to evaporation. This occurs with sharp, almost concentric boundaries, which shrink toward the centre of the tip at different rates (Plate 1, Fig. 4). The innermost boundary, enclosing the darkest region, shrinks most rapidly and the outermost, corresponding to maximum emission, least so. After the last boundary has disappeared the first layer, corresponding to enhanced emission relative to clean tungsten, remains. Its desorption occurs at slightly higher temperatures and, except in the case of neon, without a boundary.

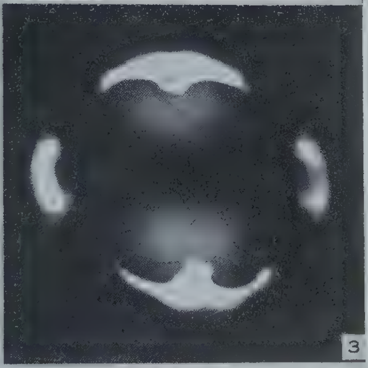
The qualitative meaning of these observations is fairly clear. At sufficiently low temperature multilayer physical adsorption occurs. The formation of the first layer is accompanied by a decrease in work function, because of the induction of a large dipole moment in each ad-atom. The resultant dipole-dipole repulsions in this layer are sufficiently strong to overcome attractive forces (except in the case of neon) and cause this layer to behave like a two-dimensional gas. The absence of repulsions with neon, where the dipole moment is small, allows attractive forces to dominate there so that even the first layer shows cohesion.

FIELD EMISSION THROUGH DIELECTRIC LAYERS









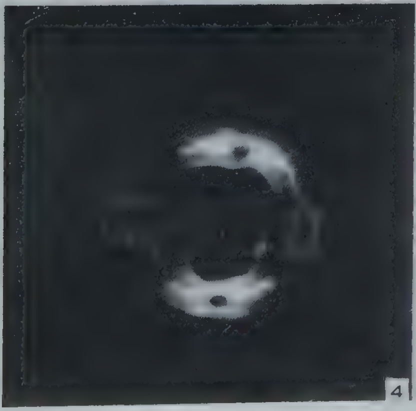


Fig. 1 (*a*).—Emission pattern from a clean tungsten tip.

Fig. 1 (*b*).—Emission pattern from a tungsten tip covered by one layer of argon.

Fig. 2.—Argon layers flowing over a tungsten tip at 21 °K. Direction of flow is from top to bottom of picture.

Fig. 3.—Emission pattern from a tungsten tip covered with five layers of argon. The multilayer cap does not extend over the entire emitting area in this photograph. The bright regions framing it correspond to a coverage of $\sim$3 layers.

Fig. 4.—Desorption of several layers of neon from a tungsten emitter. Each ring corresponds to an atomic layer.

The mechanism of dipole formation involves interaction with the metal and is necessarily of short range. Higher layers will not participate appreciably and cohesive forces will control their behaviour. They will resemble two-dimensional anisotropic liquids, with boundaries resulting from the two-dimensional analogue of surface tension. Under normal conditions evaporation will occur only from the layer edges, where atoms have the least number of nearest neighbours. Since the distance of a layer from the substrate determines its heat of desorption, the highest will evaporate most rapidly, the next one more slowly, and so on, resulting in the formation of concentric rings of different emission, with the smallest darkest one corresponding to the top layer.

TABLE 1

SUMMARY OF EMISSION DATA FOR INERT GASES ON TUNGSTEN

Values shown have been obtained from equation (2) neglecting $(\alpha_W/\alpha_G)^{2/3}$, and are based on a value of $\varphi = 4 \cdot 50$ eV for clean tungsten

Gas	φ_{mono} (eV)	Log b_W/b_{mono}	Apparent φ_{multi} (eV)	Log b_W/b_{multi}
Ne	4·35	0·8	4·05	2·2
A	3·70	1·1	4·30	1·5
Kr	3·35	1·5	3·40	2·0
Xe	3·20	0·2	3·00	0·5

TABLE 2

PHYSICAL CONSTANTS AND PERTINENT DATA FOR INERT GASES

Gas	Density* (g/cm³)	K†	$\alpha_p \times 10^{24}$† (cm³)	Crystal Radius (Å)	H_{ads}† for 1st Layer‡ (kcal)
Ne	$1 \cdot 2_{27}$	1·23	0·392	1·59	?
A	$1 \cdot 65_{40}$	1·55	1·65	1·91	1·9
Kr	$2 \cdot 1_{120}$	1·56	2·50	2·01	5·9
Xe	$2 \cdot 7_{133}$	1·82	4·10	2·20	10·0

* The subscripts in column 2 refer to the absolute temperature of the density values.

† K, dielectric constant; α_p, polarizability; H_{ads}, heat of absorption.

‡ Values obtained by the author (1958, 1959, and previously unpublished results).

The liquid-like flowing of layer boundaries is observed in all cases well below the bulk melting point of the adsorbate. This can probably be explained as follows. The lattice structure of the substrate is transmitted to the adsorbate layer by layer, although with a certain attenuation. Up to a point this prevents the latter from assuming its own bulk structure over any but the smallest distances and thus prevents long-range order. The layers accordingly behave like liquids but to different degrees. Beyond a critical thickness (which depends on temperature but, as indicated by direct counting in desorption, seems to be roughly

5–6 layers for most of the inert gases at 4–20 °K) the decrease in energy resulting from adsorbate-substrate interaction is too small to compensate for the increase caused by the lack of order in the adsorbate. Under these conditions the film becomes metastable with respect to crystallization. There is evidence that this occurs and that the crystallites can be melted by raising the temperature (Gomer 1959).

The apparent work function changes associated with mono- and multilayer adsorption are summarized in Table 1. These have been calculated by setting the factor $(\alpha_W/\alpha_G)^{2/3}$ in equation (**2**) equal to unity. Pertinent physical constants and the heats of desorption for the monolayers are given in Table 2.

IV. Discussion

(*a*) *The First Layer*

It is seen from Table 1 that work function decreases and emission increases are associated with the adsorption of one layer of inert gas. The effect was first observed by J. C. P. Mignolet (1955) with a zero field method, and must therefore be associated with the production of a dipole in the ad-complex. There are (at least) two plausible mechanisms for this. Ehrlich (1959), in agreement with Mignolet's original hypothesis, assumes the polarization of ad-atoms by the surface field of the metal. This writer favours the formation of charge transfer complexes as the cause of the dipole moments, because the large size of the adsorbate and the short range of the surface field make it impossible for the latter to " seize " more than a small fraction of each atom.

In any case the phenomenon cannot depend specifically on the presence of an applied external field, as already pointed out.

(*b*) *Field Emission through Multilayers*

The data of Table 1 indicate an apparent work function decrease for neon, an increase for argon, and essentially no changes for krypton and xenon relative to the monolayer values. In all cases, however, the total emission is decreased because of the changes in the pre-exponential terms.

It is the principal aim of this paper to find a mechanism capable of explaining these observations.

(*c*) *Mechanism*

The adsorbed film is 20–30 Å thick at maximum coverage (5–6 layers) and thus exceeds the length of the potential barrier for high energy electrons, even if allowance for the reduction in field by $1/K$, the dielectric constant, is made (Table 3). Since electrons near the Fermi level, i.e. with total energy $-\varphi$, relative to field free vacuum, contribute most heavily to field emission, the barrier region of interest is filled with inert gas atoms. Only the first layer of these contains strong dipoles so that field emission occurs through a shell of essentially non-polar dielectric.

It might seem attractive to consider the latter as an insulator and to treat the emission process as tunnelling into its conduction band. However, a perhaps more direct approach has been chosen here. It will be assumed that emission occurs by tunnelling through the potential barrier corresponding to monolayer conditions, with the following modifications.

(1) The field within the dielectric shell is reduced by $1/K$. It can be shown from elementary considerations that this will be the case if its thickness is small compared to the tip radius.

(2) The image potential is reduced by $1/K$.

(3) The potential of electrons in the shell is reduced by an average polarization energy.

(4) Each inert gas atom represents a short-range but deep attractive potential for electrons, because of the incomplete screening of the nucleus at small distances. The barrier opacity is reduced by the presence of these potential holes.

TABLE 3

CORRECTED WORK FUNCTION DATA FOR INERT GASES ON TUNGSTEN*

Gas	φ_{mono} (eV)	φ_M (eV)	α_M	$\Delta\varphi_{\text{pol}}$ (eV)	F_{vac}/K (V/Å)	$K\varphi'_{\text{mono}}/F$ (Å)
Ne	4·39	3·54	0·82	0·05	0·235	18·7
A	3·77	3·17	0·84	0·10	0·188	20·1
Kr	3·48	2·51	0·82	0·08	0·133	26·2
Xe	3·33	2·08	0·815	0·08	0·090	37·0

* φ_{mono} and φ_M refer to mono- and multilayer work functions respectively ; $\Delta\varphi_{\text{pol}}$, polarization energy ; $\varphi'_{\text{mono}} = \varphi_{\text{mono}} - \Delta\varphi_{\text{pol}}$; α_M, Nordheim image correction term ; the last column lists the length of the barrier for electrons of energy $-\varphi$.

The first two effects will tend to raise the apparent work function while the last two will reduce it. The qualitative correctness of this picture is indicated by the data of Table 1, which reflect the competition between dielectric and hole effects. In the case of neon, which has a low dielectric constant, the hole effect predominates, in argon the opposite is the case, and in krypton and xenon the effects more or less balance each other.

Effect (4) is certainly the most interesting. In order to assess it, the others must first be evaluated. This can be done very easily for (1) and (2). The presence of dielectric modifies the Fowler-Nordheim equation as follows :

$$I/V^2 = b' \exp -6 \cdot 8 \times 10^7 K \varphi^{3/2} \alpha(K)/kV, \qquad \text{........ (3)}$$

so that the effective multilayer work function φ_M is given by

$$\varphi_M = \varphi_{\text{app}} K^{-2/3} (\alpha_W/\alpha_M)^{2/3}, \qquad \text{.................. (4)}$$

where φ_{app} is the apparent work function listed in Table 1 and the α's are the appropriate Nordheim image correction factors. It can be shown readily that the argument y of α is modified by $1/K$ in the presence of dielectric :

$$y = \frac{3 \cdot 8 \times 10^{-4} F_{\text{vac}}^{\frac{1}{2}}}{K\varphi}. \qquad \text{.................. (5)}$$

φ_M can thus be determined by iteration after α_W, k, and F_{vac} have been similarly determined from the data for clean W. The corrected multilayer work functions are listed in Table 3.

(*d*) *Polarization Energy*

The eventual aim of this work will be a comparison of the atomic hole parameters with those obtained from low energy electron scattering in gases. In the latter case polarization effects at small atom-electron separations have either been ignored or lumped in with the atomic potential. In the same way we shall consider an average polarization correction valid at large electron-atom distances and shall lump effects occurring at small distances with the hole potentials.

The polarization energy will then be the "lattice" sum of the terms $\frac{1}{2}\alpha_p F_{\text{eff}}^2$ where α_p is the polarizability and F_{eff} the effective field due to the presence of an electron in the dielectric region. The latter is

$$F_{\text{eff}}=\frac{e}{Kr^2}\left(\frac{1}{1-(4/3)\pi n\alpha_p}\right), \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

r being the atom-electron distance and n the number of atoms/cm^3. The summation results in an energy term $\Delta\varphi_{\text{pol}}$

$$\begin{aligned}\Delta\varphi_{\text{pol}}&=\frac{\alpha_p e^2}{2K^2\{1-(4/3)\pi n\alpha_p\}^2}\Sigma r^{-4}\\&=\frac{7\cdot25\alpha_p N^{4/3}}{K^2\{1-2\cdot52\alpha_p N\}^2}\ \text{eV}. \qquad \ldots\ldots\ldots\ldots\ldots \quad (7)\end{aligned}$$

The quantity Σr^{-4} has been taken (Fermi 1934; Wick 1934) as $\sim 20n^{4/3}$. N is the number of moles/cm^3. The results are shown in Table 3, and can be seen to be relatively unimportant.

(*e*) *Hole Contribution*

We can now evaluate the contribution of the potential holes to the decrease in work function, $\Delta\varphi_{\text{hole}}$

$$|\Delta\varphi_{\text{hole}}|=\varphi_{\text{mono}}-\Delta\varphi_{\text{pol}}-\varphi_M. \qquad \ldots\ldots\ldots\ldots \quad (8)$$

The results are shown in Table 4.

TABLE 4

SUMMARY OF HOLE PARAMETERS

Gas	$\Delta\varphi_{\text{hole}}$ (eV)	Scattering Length, a (Å)	Hole Radius, r_0 (Å)
Ne	0·80	0·46	0·60
A	0·50	0·42	0·52
Kr	0·89	1·23	1·00
Xe	1·17	1·95	1·38

Equation (8) is based on the assumption that the contact potential associated with the first layer is unaffected by subsequent adsorption and that its origin is fundamentally different from the causes responsible for the multilayer effects. These assumptions seem justified for the following reasons. First, the monolayer contact potentials are in fairly good agreement with the zero field values after correction for image effects. Second, the emission anisotropy for multilayers is that corresponding to the monolayer. Thus the work function decrease associated with the first layer seems unaffected by subsequent adsorption and does not seem to result to any appreciable extent from the hole contributions of its atoms. This is not too surprising, even though this is important in multilayer emission, where the decreased field puts a premium on any counterbalancing effects.

We shall now endeavour to interpret $\Delta\varphi_{\text{hole}}$ in terms of the hole parameters.

(*f*) *Fermi Scattering Approximation*

The first approach utilizes Fermi's (1934) treatment of small holes in a slowly varying potential, originally developed by him to account for the term shifts of alkali atom spectra in the presence of inert gases. The method substitutes an average wave function $\bar{\psi}$ for the actual one and leads to a wave equation of the perturbed problem, identical in form to that of the unperturbed one, if the total energy of the system is modified by a term

$$\Delta W = 2\pi\hbar^2 n\mathbf{a}/m, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (9)$$

where **a** is the Fermi scattering length. The treatment is valid only for *s*-states i.e. low energy, and also breaks down unless the unperturbed potential and $\bar{\psi}$ vary slowly over distances larger than the hole dimensions. Since $\bar{\psi}$ is a relatively slowly decaying exponential here and since the hole dimensions (determined by another approximation) turn out to be small, the use of the treatment is probably justified.

Equation (**9**) implies that the solution of the perturbed problem takes the same form as that of the unperturbed one if the energy is suitably modified. In the present case this means that the Fowler-Nordheim equation may be taken over unchanged if the electron energy is increased by ΔW or, equivalently, if the work function is decreased by this amount. Thus ΔW of equation (**9**) can be equated to $\Delta\varphi_{\text{hole}}$ and the latter may be converted into a scattering length. These are shown in Table 4 and are seen to be of the expected magnitudes. The result gains a certain piquancy from the fact that it represents the outcome of a scattering experiment involving electrons of negative energy with respect to the zero of the scattering potentials.

(*g*) *Pseudo-one-dimensional Treatment*

The results of the last section taken by themselves do not provide a one to one correlation with the hole parameters. An attempt is made here to determine the hole radii directly by means of an approximation which reduces the problem to a one-dimensional one and permits the direct use of the Wentzel, Kramers, and Brillouin integral.

To do this the barrier is first sliced into (for the moment, arbitrary) elements, with the cuts orthogonal to the metal surface. The probability density of an electron parallel to the latter is constant on the metal (cathode) side of the barrier, so that the amplitude in front of each barrier element is proportional to its base area. It is now assumed that the total probability amplitude on the anode side of the barrier can be obtained by summing the anode side amplitudes for each element. It is further assumed that the latter are given by the corresponding cathode side amplitudes, multiplied by the appropriate tunnelling coefficient D_i. That is,

$$D=\sum_i (s_i/s)D_i, \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad \mathbf{(10)}$$

where s is the total and s_i the base area of the ith element.

The elements are now chosen to be prisms with base areas equal to the hole cross sections. If the inert gas layers are taken to be crystalline for these purposes there will be only a limited number of different elements. Furthermore, only the type containing the largest number of holes will contribute appreciably because of the strong exponential dependence of transmission on barrier opacity. The other elements will show up in equation (**10**) with zero weight and will seem to reduce the effective emitting area. This justifies to some extent the assumption of fixed phases implicit in this approximation.

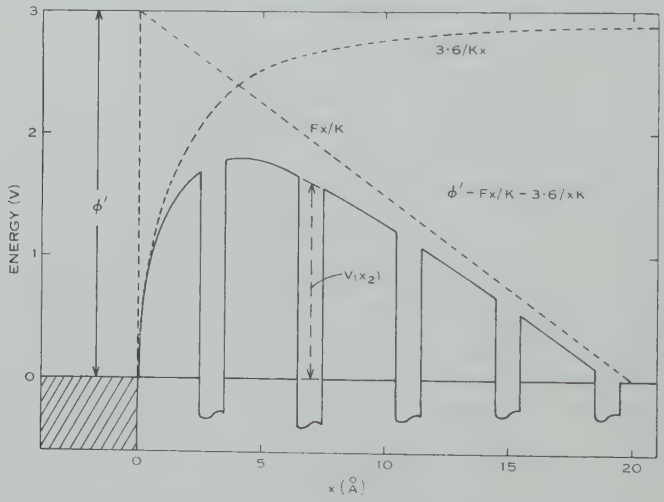


Fig. 1.—Potential diagram for the cross section of a barrier slice containing holes.

If the holes are taken to be " square " square wells of width $2r_0$ the problem becomes one-dimensional. The penetration coefficient for a barrier slice containing holes can then be found by subtracting the potential profiles of the latter from the W.K.B. integral for the solid slice. This is tantamount to the assumption that the wave function in each hole is an unattenuated oscillatory one, which so to speak transfers the decaying exponential entering the hole from the cathode side unaltered to the exit side.

The deviation of the hole potential from a square well is relatively unimportant here, as long as its bottom lies below the Fermi level, i.e. as long as the hole cuts completely through the barrier profile (Fig. 1). The detailed hole shape becomes important only if an attempt is made to calculate the reflection of the wave function at each hole, assumed to be zero in the present approximation.

The actual calculation is quite simple. The B.K.W. integral for the barrier penetration of an electron with total energy $-\varphi$ (relative to field free vacuum) can be expressed as an area A given by

$$A=\int_0^l \sqrt{(V-E)}\mathrm{d}x=\int_0^{K\varphi/F}(\varphi-Fx/K-3\cdot 6\times 10^{-8}/Kx)^{\frac{1}{2}}\mathrm{d}x \quad \ldots\ldots \quad (11)$$

$$=\tfrac{2}{3}K\varphi^{3/2}\alpha/F.$$

The appropriate A for the hole containing slice is thus

$$A=\tfrac{2}{3}K\alpha\varphi'^{3/2}_{\mathrm{mono}}/F_{\mathrm{vac}}-2r_0\sum_j V_j^{\frac{1}{2}}(x_j), \quad \ldots\ldots\ldots\ldots \quad (12)$$

where

$$\varphi'_{\mathrm{mono}}=\varphi_{\mathrm{mono}}-\Delta\varphi_{\mathrm{pol}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (13)$$

and $V(x_j)$ is the barrier height at the mid point x_j of the jth hole:

$$V(x_j)=\varphi'_{\mathrm{mono}}-Fx_j/K-3\cdot 6\times 10^{-8}/Kx_j. \quad \ldots\ldots\ldots\ldots \quad (14)$$

When the right side of equation (**12**) is equated with the corresponding expression in terms of the equivalent φ_M, $\frac{2}{3}\alpha K\varphi_M^{3/2}/F$, the following equation for r_0 results

$$r_0=\frac{K\alpha(\varphi'^{3/2}_{\mathrm{mono}}-\varphi_M^{3/2})}{3F_{\mathrm{vac}}\sum_j V^{\frac{1}{2}}(x_j)}. \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (15)$$

The values of x_j to be used in equations (**14**) and (**15**) are found from the crystal radii of the inert gases and have been chosen for the separation between 100 planes. x_1 has been chosen to be

$$x_1=\tfrac{1}{2}(d_W+d_G), \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (16)$$

where d_W and d_G refer to the crystal diameters of tungsten and the gases respectively.

The results of the calculations are shown in Table 4. The effective r_0's are seen to be considerably less than the corresponding crystal dimensions. The calculations are only slightly altered for other reasonable hole spacings (not shown in Table 4). Consequently the use of Fermi's treatment seems justified. In the present calculation the zigzagging of electrons in order to sample more holes can also be neglected, since the ratio r_0/r_{crystal} is sufficiently small to make the total path under the barrier unprofitably long in such a process.

(*h*) *Matching of the Approximations*

It is interesting to compare the results of the two approaches used. This will be done by taking the holes as square wells of depth

$$V=Ze^2/r_0, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (17)$$

Z being the effective nuclear charge. The matching of the function $u(r)$

$$u(r)=\psi/r$$

and its derivative at the hole edge $(r=r_0)$ leads to a relation between scattering length and hole strength:

$$\mathbf{a}/r_0+1=(\tan\gamma)/\gamma, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (18)$$

where

$$\gamma=(2Zr_0/r_H)^{\frac{1}{2}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (19)$$

and r_H is the first Bohr radius of hydrogen. γ is a measure of hole strength and is related to the quantity β occurring in the work of Allis and Morse (1931) by

$$\tfrac{1}{2}\gamma=\beta. \qquad \textbf{(20)}$$

The value of Z to be used in γ depends on the penetration of the scattered electron and thus on the effective r_0. Equations (**18**) and (**19**) will therefore be used to obtain Z from the previously found values of **a** and r_0. These will then be compared with the results of Allis and Morse (1931) on the scattering of slow electrons in inert gases.

The present determination of Z is not completely unequivocal since equation (**18**) has repeated roots, so that

$$Z=(r_H/2r_0)(\gamma_0+m\pi)^2, \qquad \textbf{(21)}$$

γ_0 being the smallest root of equation (**18**). However, it is possible to make the correct choices by using in equation (**20**) the values of γ lying closest to the corresponding values of 2β of Allis and Morse (1931). In some cases the choices are so obvious that this procedure is unnecessary (Table 5).

TABLE 5

EFFECTIVE Z CALCULATIONS

The numbers listed represent some of the solutions of equation (**21**). The numbers in bold type are the most appropriate choices (see text)

Gas	Z_{eff}
Ne	0·5, **7·9**, 24, . . .
A	0·6, **9·1**, 28, . . .
Kr	0·4, 5·1, **15**, 30, . . .
Xe	0·3, 3·7, 11, **22**, 37, 55, . . .

Table 6 summarizes the various results and compares them with those of Allis and Morse (1931). These authors used a cut-off Coulomb potential in their calculations. This does not differ appreciably from a square well in the region of interest. It is seen that the agreement of the Z, r_0, and β values is

TABLE 6

COMPARISON OF PRESENT RESULTS WITH DATA OF ALLIS AND MORSE*

Gas	At. No.	Z_{eff}	Z_{AM}	β	β_{AM}	r_0 (Å)	$(r_0)_{\text{AM}}$ (Å)
Ne	10	7·9	8·3	2·06	1·71	0·60	0·37
A	18	9·1	10·6	2·1	2·7	0·52	0·74
Kr	36	15·0	15·7	3·8	3·66	1·00	0·90
Xe	54	22	[22]	5·3	[4·7]	1·38	[1·06]

* The columns subscripted AM refer to the results of Allis and Morse (1931), the others to the present work. $\beta=\frac{1}{2}\gamma$. The AM data on xenon have been found by extrapolation.

quite good, considering the relative crudity of our approximations. It can be concluded that the mechanism proposed here is essentially correct and that the approximations made in evaluating it are reasonable.

V. Acknowledgments

The author would like to thank Dr. C. S. Wright for pointing out Fermi's treatment, and many of his other colleagues for stimulating discussions. He also wishes to thank Mr. J. Dafler for carrying out many of the numerical calculations. This work was supported in part by a grant from the California Research Corporation. Its presentation at the Melbourne Conference on Solid State Physics was made possible by assistance from the United States Atomic Energy Commission.

VI. References

Allis, W. P., and Morse, P. M. (1931).—*Z. Phys.* **70** : 567.

Ehrlich, G., and Hudda, F. G. (1959).—*J. Chem. Phys.* **30** : 493.

Fermi, E. (1934).—*Nuovo Cim.* **11** : 157.

Gomer, R. (1958).—*J. Chem. Phys.* **29** : 441, 443.

Gomer, R. (1959).—*J. Phys. Chem.* **63** : 468.

Gomer, R., Wortman, R., and Lundy, R. (1957).—*J. Chem. Phys.* **26** : 1147.

Good, R. H., and Müller, E. W. (1956).—"Handbuch der Physik." Vol. 21, p. 176. (Springer-Verlag : Berlin.)

Mignolet, J. C. P. (1955).—*Rec. Trav. Chim. Pays-Bas.* **74** : 685.

Wick, G. C. (1934).—*R.C. Accad. Lincei* **19** : 27.

BOMBARDMENT OF METALS BY INERT GAS IONS

By G. J. Ogilvie*

[*Manuscript received September 25, 1959*]

Summary

When the surface of a metal crystal is bombarded by inert gas ions, metal atoms are knocked off the surface in certain preferred directions. Evidence is presented to show that some atoms are also knocked into the crystal. A thin disoriented surface layer appears which, if the crystal is an alloy, is of uniform but different composition from that of the original crystal. The importance of these findings to some current ideas about damage due to bombardment is briefly discussed and it is shown that present theories do not account satisfactorily for the observations.

I. Introduction

When a metal is bombarded by positive ions material is removed from its surface. The results of the large number of investigations of this phenomenon have been reviewed by Massey and Burhop (1952) and Wehner (1955).

In the present paper some experiments are described which have given information about the changes produced at the surfaces of crystals by bombardment with inert gas ions having energies up to 4 keV. The information was derived from electron diffraction patterns made with 50 kV electrons. Transmission diffraction patterns were obtained from parts of the specimen thinned by electropolishing. Reflection diffraction patterns were also studied, but they give relatively little information about the effects described in the next section. The apparatus and other details of the experimental technique have been described by Ogilvie (1959) and Gillam (1959).

II. Results

(*a*) *Disorientation produced by Ion Bombardment*

Single crystals of silver with (111), (100), and (110) planes parallel to the surface have been bombarded with ions of rare gases having energies between about 10 eV and 4 keV. Most of the results described below were obtained by using argon ions but essentially the same results were obtained with helium or xenon ions. Electron diffraction patterns obtained from a crystal are changed by ion bombardment: interferences appear in the photographs which are not present in patterns obtained from an electropolished crystal. These additional interferences can be indexed in terms of the silver lattice.

A pattern obtained from a crystal with (111) parallel to the surface is given in Plate 1, Figure 1, where the indices of most of the interferences are given and the extra interferences are distinguished from the original ones. Such patterns can be accounted for in detail by assuming that many small crystallites appear

---

* Division of Tribophysics, C.S.I.R.O., Melbourne.

in the region of the surface of the crystal after bombardment, the orientations of these crystallites being related to the parent orientation by tilts about axes lying in the plane of the surface of the crystal (Ogilvie 1959). These tilts may be as large as 35° but they are usually less than 15°.

Similar patterns have been obtained for ion energies down to 20 eV. Crystals of other orientations give patterns allowing the same interpretation. For crystals with (100) or (110) parallel to the surface an almost random disorientation is found for ion energies of about 10 eV. In these experiments the temperature of the crystal during bombardment did not exceed 80 °C.

A. A. Thomson and the author have shown that, as the temperature of the crystal during bombardment is increased, there appears in the diffraction patterns a progressively expanding zone, centred on the undiffracted electron beam, free from extra interferences due to tilted crystallites. This implies that the maximum tilt of the crystallites is decreasing. At 400 °C the tilts are not greater than about 3°. In patterns from a crystal having crystallites with a small angular range, a set of interferences was prominent which had been partly masked when a large angular range was present. This point is illustrated in Plate 1, Figure 2, where a regular network of relatively faint spots is shown having a geometrical relationship to the main spots of a crystal with (111) parallel to the surface. The primary spots can be indexed by comparison with Plate 1, Figure 1. The distance of these extra spots from the centre of the pattern changes as the crystal is tilted. In crystals of other orientations streaks have been found after ion bombardment under similar conditions. On the basis of the analysis of electron diffraction from crystals containing stacking faults by Whelan and Hirsch (1957*a*, 1957*b*), these interferences are thought to be due to stacking faults produced on (111) planes in the crystal caused by the bombardment. Streaks have also been found which imply that a disturbance may exist on {100} planes.

(*b*) *Bombardment of Composite Targets*

A series of experiments has been performed by A. A. Thomson and the author in which, prior to bombardment, silver crystals were coated with evaporated copper films. The thickness of the film was estimated from the weight of copper evaporated and the distance between the evaporation source and the specimen. A diffraction pattern from a crystal with (111) parallel to the surface coated with 100 Å of copper and then bombarded for a short time is given in Plate 2, Figure 1. This crystal was bombarded with 10^{17} ions/cm^2, which for a bulk specimen of copper is equivalent to the removal of approximately 3×10^{16} atoms/cm^2 (Wehner 1955). The pattern is similar in its general appearance to Plate 1, Figure 2, but through the most intense spots there are streaks bearing a geometrical relationship to the rest of the pattern. These streaks, which are consistent with faults existing on {111} planes that are not parallel to the surface of the crystal, were very weak or absent when the crystal did not have a contaminating film on it prior to bombardment. Consequently, it seems likely that the faults are due to copper atoms embedded in the crystal and there segregating on {111} planes (for example, see Geisler 1951).

(c) *Bombardment of Alloys*

When an alloy specimen is bombarded a thin layer of changed composition is formed on the surface. The phenomenon has been studied in detail by Gillam (1959). A diffraction pattern from an ordered, coarsely crystalline specimen of Cu_3Au bombarded with argon ions is shown in Plate 2, Figure 2. In this pattern the interferences are doubled with the exception of those spots which appear because of the existence of order in the alloy. The inner of each of the pairs of interferences comes from the surface layer and indicates that the surface layer is enriched in gold. After bombardments with 30 eV and 3 keV argon ions the surface layer has lattice parameters equivalent to gold contents of 68 and 45 atomic per cent. respectively. Since the interferences from the surface layer are not markedly broadened, this layer must be approximately constant in composition throughout its thickness and there must be a sudden change in composition at the interface. Although this surface layer is partly disoriented and its order has been destroyed, much of it has the same orientation as the part of the specimen immediately below.

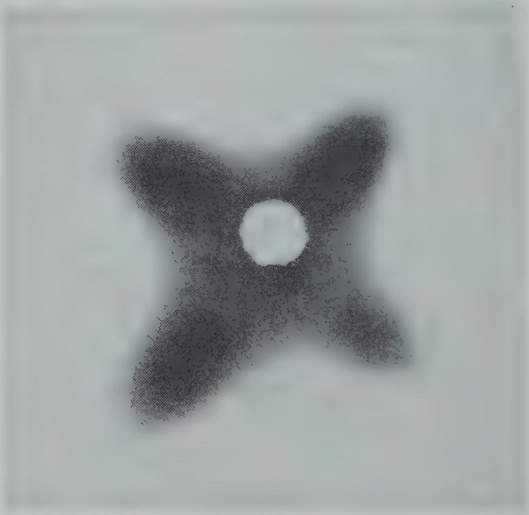


Fig. 1.—Deposit spot pattern obtained from a silver single crystal bombarded with 4 keV argon ions. An ion beam passed through the hole in the collector plate to bombard the specimen. The [100] direction of the crystal was approximately parallel to the direction of the ions.

The same effects have been found with helium and xenon ions and for bombarding energies in the range 14 eV to 3 keV. A general result is that the lower the energy of the ions, the greater is the gold concentration of the surface layer.

The thickness of the altered layer has been measured using polycrystalline specimens of Cu_3Au and evaporating a calibrating layer of copper of known thickness onto the specimen after it had been bombarded. In an electron diffraction pattern obtained from the composite specimen the intensities of the rings from the copper layer and the altered layer were compared and the thickness of the altered layer deduced from the ratio of these intensities. After bombarding with 70 eV argon and xenon ions the thickness obtained by this method was about 30 Å, while for a bombardment with 70 eV helium ions a thickness for the altered layer of about 80 Å was found. The error in these measurements is probably less than ±20 per cent.

(*d*) *Deposit Spot Patterns*

Wehner (1955, 1956) has shown that, when bombarded by mercury ions of energies between 50 and 150 eV, atoms of copper and silver single crystals tend to be ejected in ⟨110⟩ directions, while atoms of iron and tungsten tend to be ejected in ⟨111⟩ directions.

H. G. Scott and the author have studied the patterns obtained when the atoms knocked off a single crystal of silver by bombardment with inert gas ions are collected on a glass plate placed near the crystal. The collector plate was separated from the crystal by about 1 cm and the ions were directed through a hole in the glass plate onto the crystal. The patterns obtained in this way, called deposit spot patterns by Wehner, show spots centred on the intersection of the glass plate and the ⟨110⟩ directions. A pattern obtained from a crystal bombarded with argon ions of 4 keV energy is given in Figure 1. Very similar patterns have been obtained with ion energies down to a few hundred volts. Helium ions also give very similar patterns in the same energy range.

III. Discussion

Since disorientation and stacking faults seem to be produced by low energy ion bombardment (down to about 10 eV) and since it is most unlikely that a single ion impact, at these low energies at least, will produce a dislocation or a stacking fault, it becomes attractive to imagine that these defects appear as a consequence of the formation of point defects by the bombardment. The condensation of point defects to produce stacking faults has been demonstrated experimentally by Hirsch *et al.* (1958) and Silcox and Hirsch (1959), for aluminium and gold respectively. A means by which disoriented crystallites may be produced by ion bombardment has been discussed by Ogilvie (1959). The electron diffraction pattern from the silver crystal which had a copper layer on it before bombardment demonstrates that atoms on the surface have an appreciable probability of being driven into the crystal. On the other hand, the existence of a layer of altered and roughly uniform composition on the surface of an alloy specimen after bombardment implies a transport, during the bombardment, of atoms from the interface between the layer and the rest of the specimen towards the surface.

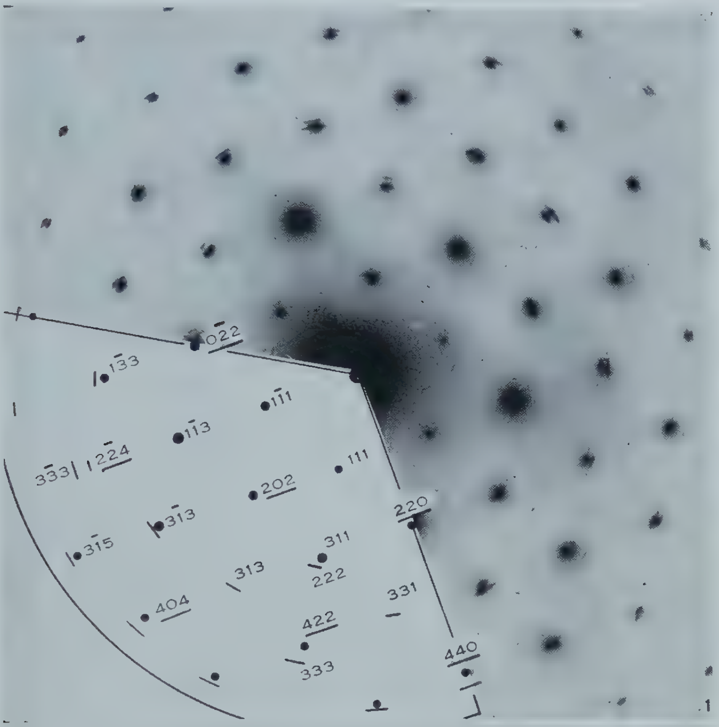
Recently, two proposals have been made to account for the ejection of atoms from a metal crystal by bombardment with heavy particles. The first (Henschke 1957*a*, 1957*b*) is that a particle incident on the crystal rebounds from an atom just below the surface and strikes an atom in the surface on the underside and hence propels it away from the crystal. This proposal may be satisfactory if only one atom is ejected per incident ion. However, Keywell (1955) has shown that an argon ion of 4 keV will eject about seven atoms from a silver crystal. Our deposit spot patterns have been obtained with ions of this energy and are very similar to those obtained by Wehner (1956). Thus, the theory provides no explanation for the directional ejection of most of the atoms. The first atom hit by the incident particle must have a large fraction of the energy of the incident particle transferred to it because, in Henschke's theory, that particle has to be deflected through a large angle in order to strike a surface atom. This atom will be displaced from its site in the crystal and it is possible that it could be deflected by other atoms of the crystal so that it escapes from the crystal. Another possibility is that, by colliding with other atoms in the crystal, this atom may transfer to the other atoms a motion in the right direction with sufficient energy to cause them to escape.

The second proposal (Wehner 1956) is that the impact between the incident particle and a surface atom produces a " shock wave " proceeding along a close-packed row of atoms and this returns to the surface to eject one or more atoms. Wehner does not specifically explain how the shock wave returns to the surface in order to eject atoms from it.

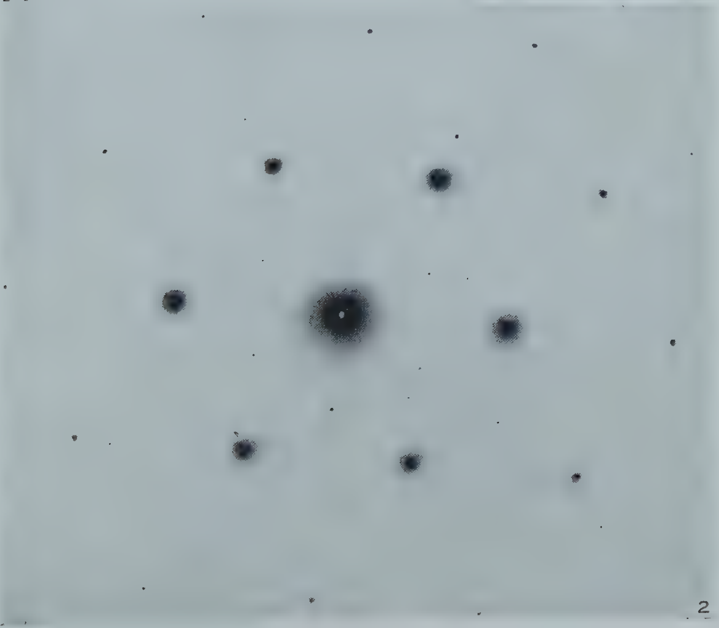
Silsbee (1957) has examined this possibility in more detail. He showed that it is proper to treat the processes transferring energy through a metal crystal as a sequence of two-body collisions if the energy transferred is greater than about 0·1 eV. He also showed that this sequence of two-body collisions may tend to be directed along the most closely packed directions in the crystal, and that an energy pulse may be propagated in this manner a considerable distance through the crystal. He showed that for pure copper the principal energy loss as the pulse propagates along a close-packed row of atoms is due to an interaction between neighbouring rows of atoms. The energy loss is 2 per cent. for a collision when 100 eV is transferred, 7 per cent. for 10 eV, and 20 per cent. for 1 eV. This mechanism may operate in pure metals but it seems not to work in the case of alloys where the efficiency of energy transfer is much reduced if the successive atoms in the chain have very different masses. In the copper-gold alloys studied, an energy loss of about 10 per cent. per collision must exist for this reason. This means that for 70 eV helium ion bombardment, in the absence of losses caused by interaction between rows of atoms in disordered Cu_3Au, a pulse could propagate on the average about 18 Å and still retain sufficient energy to displace an atom (say, 5 eV). The other losses will reduce this to less than 10 Å. However, experimentally, it was found (Gillam 1959) that for 70 eV helium ion bombardment the layer was 80 Å thick. Thus Silsbee's mechanism does not directly account for the observed thickness of the altered layer on bombarded alloy specimens.

Since disorientation of the surface of a single crystal by ion bombardment can be very marked, the intensity distribution over a deposit spot pattern obtained from that crystal is likely to be influenced. Wehner (1956) and Henschke (1957*a*, 1957*b*) have devised explanations to account for changes that Wehner found in the details of the spot patterns for ion energies of about 40 eV. These explanations do not take into account the possibility of surface disorientation and consequently should be reconsidered.

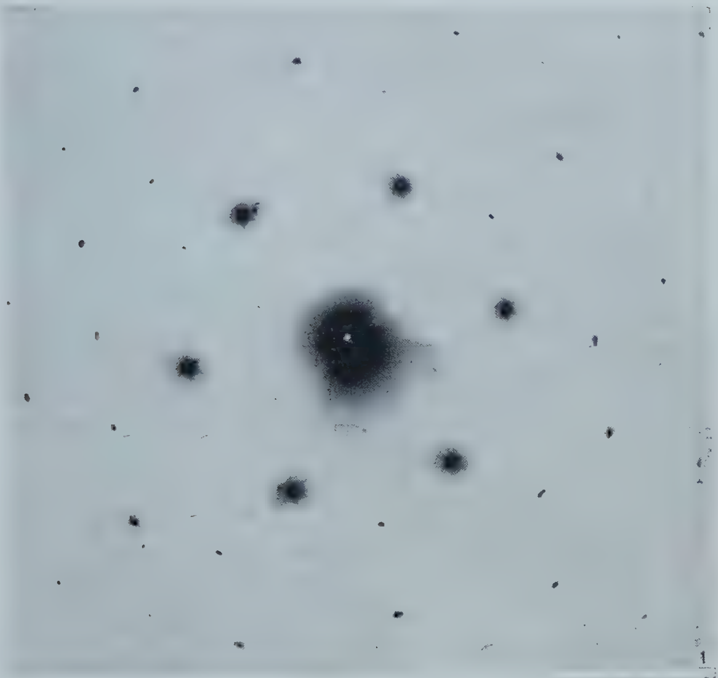
BOMBARDMENT OF METALS BY INERT GAS IONS



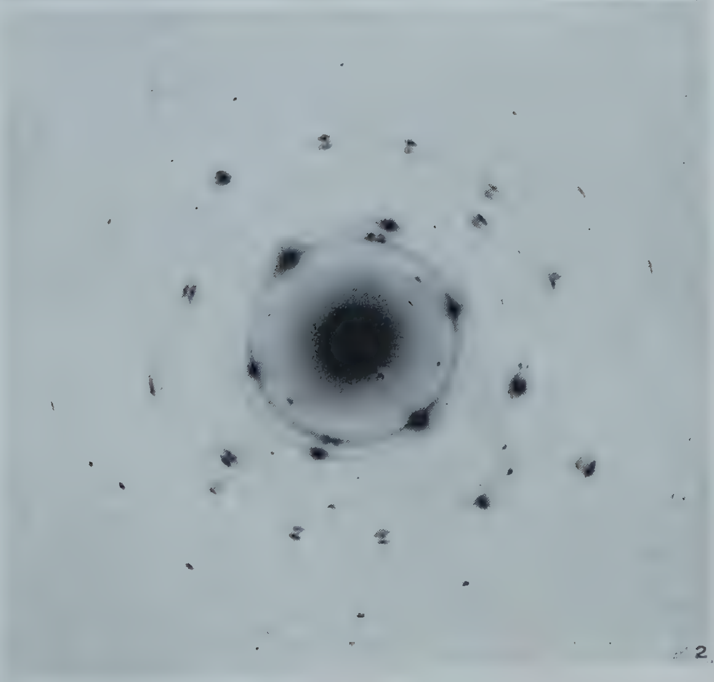
1



2



1



2

IV. ACKNOWLEDGMENTS

The author is grateful to Dr. H. G. Scott and Mr. A. A. Thomson for permission to quote experimental results before their detailed publication elsewhere.

V. REFERENCES

GEISLER, A. H. (1951).—" Phase Transformations in Solids." (Ed. R. Smoluchowski.) p. 387. (Wiley : New York.)

GILLAM, E. (1959).—*J. Phys. Chem. Solids* **11** : 55.

HENSCHKE, E. B. (1957*a*).—*J. Appl. Phys.* **28** : 411.

HENSCHKE, E. B. (1957*b*).—*Phys. Rev.* **106** : 738.

HIRSCH, P. B., SILCOX, J., SMALLMAN, R. E., and WESTMACOTT, K. H. (1958).—*Phil. Mag.* **3** : 897.

KEYWELL, F. (1955).—*Phys. Rev.* **97** : 1611.

MASSEY, H. S. W., and BURHOP, E. H. S. (1952).—" Electronic and Ionic Impact Phenomena." (Oxford Univ. Press.)

OGILVIE, G. J. (1959).—*J. Phys. Chem. Solids* **10** : 222.

SILCOX, J., and HIRSCH, P. B. (1959).—*Phil. Mag.* **4** : 72.

SILSBEE, R. H. (1957).—*J. Appl. Phys.* **28** : 1246.

WEHNER, G. K. (1955).—" Advances in Electronics and Electron Physics." Vol. 7. (Ed. L. Marton.) p. 239. (Academic Press : New York.)

WEHNER, G. K. (1956).—*Phys. Rev.* **102** : 690.

WHELAN, M. J., and HIRSCH, P. B. (1957*a*).—*Phil. Mag.* **2** : 1121.

WHELAN, M. J., and HIRSCH, P. B. (1957*b*).—*Phil. Mag.* **2** : 1303.

EXPLANATION OF PLATES 1 AND 2

PLATE 1

Fig. 1.—Transmission electron diffraction patterns from a silver crystal with surface parallel to (111) bombarded with argon ions (10^{19} ions/cm^2, 3 keV, temperature of crystal during bombardment less than 80 °C). The indices of the interferences are shown and the spots which appear when the crystal is examined before bombardment are shown by bars beneath the index symbols.

Fig. 2.—Transmission electron diffraction pattern for a silver crystal with surface parallel to (111) bombarded at an elevated temperature (10^{19} ions/cm^2, $1 \cdot 3 \times 10^{16}$ ions/cm^2 sec, 130 eV, temperature of crystal during bombardment 250 °C).

PLATE 2

Fig. 1.—Transmission electron diffraction pattern from a silver crystal with surface parallel to (111). The crystal had been coated with an evaporated copper layer 100 Å thick before bombardment with 10^{17} argon ions/cm^2 at 130 eV in one minute.

Fig. 2.—Transmission electron diffraction pattern from a coarse-grained specimen of ordered Cu_3Au after bombardment by argon ions. Most interferences are split into two components with different lattice parameters. Those spots which are not split radially correspond to superlattice reflections.

NUCLEATION AND GROWTH OF ICE CRYSTALS UPON CRYSTALLINE SUBSTRATES

By N. H. Fletcher*

[*Manuscript received August* 20, 1959]

Summary

The importance of various surface imperfections in the nucleation of ice crystals is considered and it is concluded that dislocations are not preferred nucleation sites, whereas steps and re-entrant corners do encourage nucleation. Expressions are derived for the height of the free energy barrier in several important cases, and the retention of embryos under unsaturated conditions is discussed. Embryos initially grow by two-dimensional nucleation, but ultimately a dislocation growth mechanism becomes more important. There is not yet enough experimental evidence to decide upon the reality of the entropy effect previously proposed, which inhibits nucleation upon certain crystal faces.

I. Introduction

The basis of the theory of the nucleation of new phases was established long ago by Volmer and Weber (1925) and Becker and Döring (1935), and remains fundamentally unchanged today. Within a supersaturated vapour, or a supercooled liquid, there are transient groupings of molecules having the structure of the stable phase—liquid or solid as the case may be. Because of their surface free energy these embryos are unstable and are continually being created and destroyed by thermal fluctuations, in such a way that their numbers maintain a Boltzmann distribution in energy.

The free energy barrier associated with the formation of an embryo has a maximum value at a certain critical size. Smaller embryos tend to disappear, but once an embryo exceeds this critical size it can grow without limit. The probability of an embryo growing to critical size is a very sharp function of supersaturation or supercooling, so that the nucleation process has a sharp threshold which is relatively easy to determine experimentally.

Homogeneous nucleation within the pure substance itself presents a large free energy barrier, and requires supersaturations of hundreds of per cent. or supercoolings of many tens or even hundreds of degrees. The presence of foreign particles or surfaces lowers the barrier greatly and most phase transformations take place by such a process of heterogeneous nucleation.

The theory of homogeneous nucleation is by now fairly well developed, though on a molecular basis many of its concepts are rather unsatisfactory. Similarly, nucleation of crystals upon perfect crystalline substrates (Turnbull and Vonnegut 1952) and upon perfect spherical particles (Fletcher 1958) have been dealt with, at least to a first approximation. The purpose of the present paper is to discuss the role of surface imperfections in the substrate upon its efficiency as a nucleation catalyst. The particular case of conical pits in an otherwise flat surface has been considered by Turnbull (1950) on a non-crystalline approximation, but here we shall be concerned with such specifically crystalline structures as growth steps and dislocations.

---

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.: present address: Department of Physics, University of New England, Armidale, N.S.W.

II. Nucleation Theory

Before we consider the nucleation and growth of a crystal embryo on a substrate, we must make some detailed assumptions about the properties of this embryo, and in particular about its surface. The structure of a liquid surface is disordered. so that all surface orientations have the same free energy per unit area, and surface free energy will generally be minimized when the surface assumes the shape of a part of a sphere. The surface free energy of a crystal face, on the other hand, is strongly dependent upon the orientation of the face relative to the crystal axes. The surface free energy of a crystal will therefore generally be a minimum when the crystal has some typical polyhedral habit. Burton and Cabrera (1949) have discussed the structure of such habit faces in some detail and their results indicate that for many materials near room temperature these faces are, in equilibrium, almost atomically flat. Higher index faces, however, or even habit faces at higher temperatures may become rough on an atomic scale.

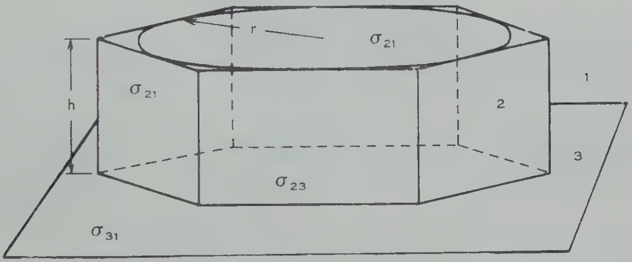


Fig. 1.—A prismatic embryo growing upon a perfect plane substrate. Subscript 1 refers to the parent phase, 2 to the embryo, and 3 to the substrate.

A crystalline embryo growing on a smooth substrate thus has the general appearance shown in Figure 1, which is drawn for the case of a crystal whose habit is a hexagonal prism. The geometry of such a prismatic embryo is characterized by its height h, the inscribed radius r, and the number n of prism faces.

If we define a parameter ξ by

$$\xi=(n/\pi)\tan(\pi/n), \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

then the free energy of formation of the embryo is

$$\begin{aligned}\Delta G&=(\text{volume free energy})+(\text{surface free energy})\\&=\pi r^2\xi h\Delta G_v+\pi r^2\xi(\sigma_{23}-\sigma_{31})+(\pi r^2\xi+2\pi r\xi h)\sigma_{21}, \qquad \ldots\ldots \quad (2)\end{aligned}$$

where ΔG_v is the free energy difference per unit volume of crystal between the parent phase and the new bulk crystalline phase, and σ is the interface free energy per unit area. Subscripts 1, 2, and 3 refer to parent phase (vapour or liquid), crystalline embryo phase, and substrate respectively. For the sake of simplicity we have assumed all habit faces of the embryo to have the same value of σ_{12}.

If we define a quantity γ, which essentially measures the disparity between the embryo and substrate, by

$$\gamma=(\sigma_{21}-\sigma_{31}+\sigma_{23})/\sigma_{21}, \quad \ldots\ldots\ldots\ldots \quad (3)$$

then (2) can be simplified to

$$\Delta G=\pi r^2\xi h\Delta G_v+\pi r^2\xi\gamma\sigma_{21}+2\pi r\xi h\sigma_{21}, \quad \ldots\ldots\ldots \quad (4)$$

and we shall henceforth simply use σ to mean σ_{21}.

In a supercooled or supersaturated parent phase ΔG_v is negative, while $0 \leqslant \gamma \leqslant 2$. The value of ΔG initially increases with increasing r and h, but decreases again after a critical size is reached. To find this critical condition we require

$$\partial\Delta G/\partial h=0, \quad \partial\Delta G/\partial r=0. \quad \ldots\ldots\ldots\ldots \quad (5)$$

The first of these gives the critical radius

$$r^*=2\sigma/\Delta G_v, \quad \ldots\ldots\ldots\ldots \quad (6)$$

while the second gives the equilibrium habit

$$h=r\gamma. \quad \ldots\ldots\ldots\ldots \quad (7)$$

This second result could also have been obtained by use of Wulff's rule (Wulff 1901), which states that in equilibrium the perpendicular distance of any face from the centre of a crystal is proportional to its free energy per unit area.

Substitution of (6) and (7) into (4) leads to the result

$$\Delta G^* = 4\pi\xi\sigma^3\gamma/\Delta G_v^2. \quad \ldots\ldots\ldots\ldots \quad (8)$$

Provided that the number of molecules associated into embryos is much less than the total number of molecules of parent phase in contact with the substrate, the number of critical embryos per unit area is given by Boltzmann's distribution law as

$$n \exp(-\Delta G^*/kT), \quad \ldots\ldots\ldots\ldots \quad (9)$$

where n is the number of parent phase molecules in contact with unit area of substrate. The rate at which critical embryos gain an additional molecule to become freely growing is the nucleation rate J and is given by

$$J=An \exp(-\Delta G^*/kT). \quad \ldots\ldots\ldots \quad (10)$$

The kinetic constant A depends upon details of the system considered.†

---

† Strictly speaking J should be an integral taking into account all possible ways of crossing the free energy barrier of which ΔG^* is the height of the saddle point. As far as the author is aware no detailed consideration has yet been given to this point, and all theories use a simple one-dimensional approximation as we have done here. It is possible that this consideration could make an order of magnitude difference to the rate J. In view of similar uncertainties in the quantity A, however, this is not of any practical importance in the present state of the theory.

For nucleation from a supersaturated vapour

$$A \sim 4\pi r^{*2} p/(2\pi mkT)^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{11})$$

(Volmer and Weber 1925), where p is the partial pressure of the vapour and m the mass of a vapour molecule. A more exact expression for A has been given by Becker and Döring (1935) but this will not concern us here. The value of ΔG_v in this case can be written explicitly as

$$\Delta G_v = (-kT/V) \ln (p/p_\infty), \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{12})$$

where V is the volume occupied by a molecule in the crystal, and p_∞ is the vapour pressure in equilibrium with a crystal of infinite extent.

For nucleation from a liquid the kinetic coefficient has been derived by Turnbull and Fisher (1949) to be

$$A \sim (kT/h) \exp (-\varphi/kT), \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{13})$$

where φ is the activation energy for short-range diffusion in the liquid. In this case ΔG_v is given approximately by

$$\Delta G_v = -|\Delta S_v|\Delta T, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{14})$$

where ΔS_v is the entropy of fusion per unit volume of crystal and ΔT is the supercooling of the liquid.

In either case, for substances with fairly average properties the kinetic coefficient A is of order 10^{25} cm^{-2} sec^{-1}, an uncertainty of a few orders of magnitude having little effect since the behaviour of J is completely dominated by the exponential term in (**10**).

We have not, as yet, taken any account of the molecular nature of the crystal and substrate. This has been discussed by Turnbull and Vonnegut (1952), who show that two effects must be considered. In the first place, the interfacial free energy σ_{23} will depend upon the dislocation content of the interface, which in turn depends upon the relation between the crystal structures of embryo and substrate. This effect contributes to the value of γ and is automatically taken into account in our discussion. If, however, the structure and cell dimensions of the surfaces of substrate and embryo are very similar the interface may contain no dislocations, the misfit being taken up in elastic strain ε in the embryo. In this case a strain energy term of order $C\varepsilon^2$ must be added to ΔG_v as given by (**12**) or (**14**). This added term means that a larger supersaturation or supercooling is required for nucleation than would be the case if the embryo were unstrained. We shall mention this effect again later.

III. Nucleation on Dislocations

The first topic to be discussed is the role of dislocations as nucleation sites. Edge dislocations can be dismissed almost immediately, as they will, except in very special circumstances, merely contribute more dislocations to the interface, or elastic strain to the embryo. The case to be considered is that of a screw dislocation meeting an otherwise perfect face. An embryo growing on such a dislocation will suffer by reason of elastic strain, but will be able to add further molecular layers by the Frank mechanism (Frank 1949) instead of requiring surface nucleation. The question to be considered is whether the combined effect raises or lowers the effective ΔG^*.

The two possible situations are shown in Figure 2 where polygons have been replaced by circles for clarity. Figure 2 (*a*) shows a prismatic embryo with a single molecular layer or two-dimensional embryo growing on its upper surface, while Figure 2 (*b*) shows a dislocated embryo. It must be admitted that these pictures are not complete since they provide no mechanism for lateral growth of the embryo, but they do represent a closer approach to reality than the perfect embryo of Figure 1. It should be emphasized that the dimensions R and r are not necessarily equal in the two figures.

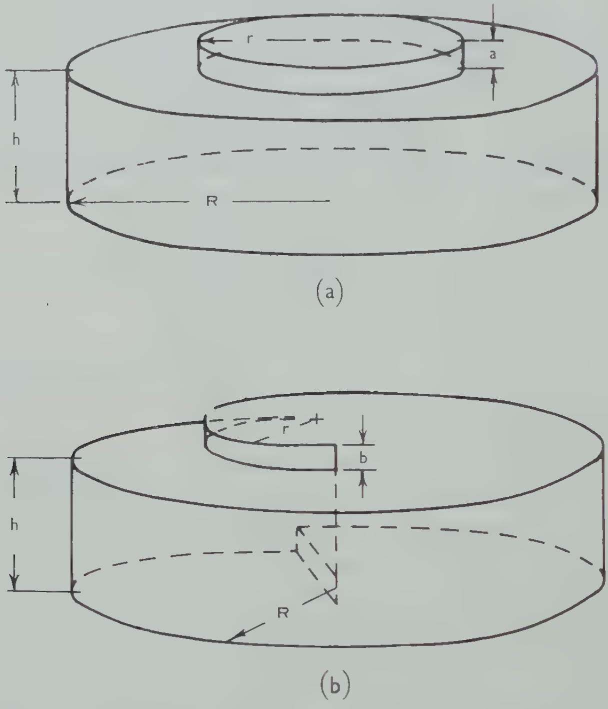


Fig. 2.—Crystalline embryos growing upon (*a*) a perfect substrate, and (*b*) a substrate from which emerges a screw dislocation with Burgers vector *b*.

The free energy of formation of these two structures can be written down analogously to equation (**2**). For the non-dislocated embryo we have

$$\Delta G_1 = \pi R^2 \xi h \Delta G_v + \pi R^2 \xi \gamma \sigma + 2\pi R \xi h \sigma + \pi r^2 \xi a \Delta G_v + 2\pi r \xi a \sigma, \quad \ldots\ldots\ldots\ldots \quad (\mathbf{15})$$

where a is the unit cell dimension along the axis normal to the substrate. For the dislocated embryo

$$\Delta G_2 = \pi R^2 \xi h \Delta G_v' + \pi R^2 \xi \gamma \sigma + 2\pi R \xi h \sigma + \tfrac{1}{2}\pi r^2 \xi \Delta G_v' + \pi r \xi b \sigma + 2rb(\gamma - 1)\sigma, \quad (\mathbf{16})$$

where $\Delta G_v'$ includes the contribution of elastic strain energy and b is the Burgers vector of the dislocation.

When the critical values ΔG_1^* and ΔG_2^* are evaluated as before it is found that

$$r^* = \tfrac{1}{2} R^* \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{17})$$

as drawn, and that

$$\Delta G_1^* = \frac{4\pi\xi\sigma^3\gamma}{\Delta G_v^2} - \frac{\pi a\xi\sigma^2}{\Delta G_v}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (18)$$

and

$$\Delta G_2^* = \frac{4\pi\xi\sigma^3\gamma}{\Delta G_v'^2} - \frac{\pi a\xi\sigma^2}{\Delta G_v'}\left(2\gamma - \frac{3}{2}\right). \quad \ldots\ldots\ldots \quad (19)$$

The second term in (**18**) is positive, whilst the second term in (**19**) is negative for small γ. However, since $|\Delta G_v'| < |\Delta G_v|$ the first term of (**19**) is larger than the corresponding term in (**18**) by an amount which can be found for specific cases.

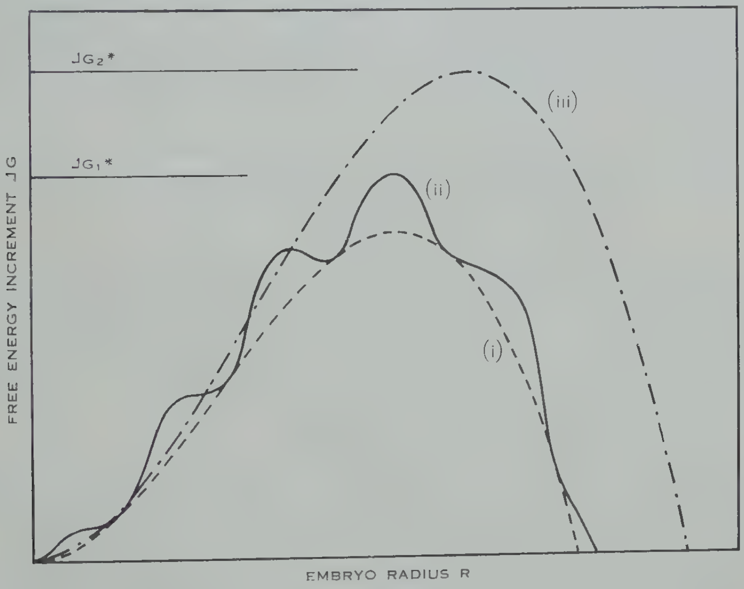


Fig. 3.—Free energy barrier to nucleation under various assumptions: (i) simple embryo with surface nucleation neglected; (ii) simple embryo including surface nucleation, (iii) embryo containing a screw dislocation.

The situation is illustrated in Figure 3 where we have drawn curves of ΔG as a function of R for the simple case considered in Section II, for an embryo with surface nucleation, and for an embryo containing a screw dislocation. Whilst the dislocated embryo is presented with a simple nucleation barrier, the barrier for an embryo requiring surface nucleation is multiple. Such multiple barriers have been discussed by Giddings and Eyring (1958) who show that only barriers lying within about kT of ΔG^* need be considered, the kinetic coefficient A in (**10**) being divided by the number of barriers in this range. Typically $\Delta G^* \approx 60kT$ in this sort of nucleation process, and the total number of barriers (i.e. of molecular layers in the embryo) is usually of the order of ten. We can therefore neglect the multi-barrier aspect of the situation and consider simply the maximum height of the barrier.

Frank (1951) has considered the strain energy associated with screw dislocations in another connexion and we can use some of his results together with Penny's values for the elastic constants of ice (Penny 1948) to calculate this case. Frank gives the strain energy density as

$$\left.\begin{array}{ll} \mu b^2/8\pi r^2 & (r \geqslant 5b/\pi), \\ \mu b/20\pi r - \mu/200 & (r \leqslant 5b/\pi), \end{array}\right\} \qquad \ldots\ldots\ldots\ (20)$$

where μ is the modulus of rigidity, and we shall take the Burgers vector b equal to the lattice constant along the z-axis.

The result for ice is that nucleation upon dislocations is impossible at temperatures above about -2 °C, and even at -20 °C dislocations are favoured nucleation sites only if $\gamma < 0 \cdot 005$. The situation is likely to be similar in most other materials, since according to Frank σ/μ does not vary widely, and it is essentially this ratio which determines the relative importance of surface and elastic energies.

We may thus dismiss dislocations as favoured nucleation sites, and in the next section we shall consider some other types of surface imperfections.

IV. NUCLEATION IN STEPS AND CORNERS

The nucleation of crystals in steps or re-entrant corners of a substrate is most simply considered for the case of a hypothetical simple cubic crystal in which interactions are with nearest neighbours only. For such a crystal only {100} habit faces will even exist under equilibrium conditions, a {110} face, for example, being stepped on an atomic scale and having a surface free energy per unit area equal to $\sqrt{2}$ times that of a {100} face.

We shall consider only right-angled steps and corners, as shown in Figure 4, the contact constant being γ on all interfaces. The theory is developed exactly as in Section II of this paper, except that it is convenient to work in terms of the lengths of the three independent edges of the embryo, rather than in terms of an inscribed radius. The result in this case is particularly simple, the critical free energy being

$$\Delta G^* = \frac{32\,\sigma^3}{\Delta G_v^2}\left(\frac{\gamma}{2}\right)^n, \qquad \ldots\ldots\ldots\ (21)$$

where n is the number of planes with which the embryo is in contact ($n=0$, 1, 2, or 3).

Since for almost any surface $\gamma < 2$, it is clear that corners ($n=3$) and steps ($n=2$) provide sites at which crystals can nucleate much more readily than upon a flat surface.

Equation (**21**) actually represents a simplification in that the two-dimensional nucleation which we considered in Section III has been neglected. This involves neglect of a term like the second in equation (**18**) and is not serious unless γ is small. When γ becomes zero, however, representing nucleation upon a substrate of the same material, the effect of this neglected term is important, as is manifest in the theory of crystal growth.

Another interesting point is that even in a corner and for $\gamma \ll 1$ there is a finite barrier to nucleation, and a finite supersaturation or supercooling is required to form a stable embryo. This is a consequence of the regularity of the crystal lattice and the assumption of only nearest neighbour interactions. These work to preclude the formation of non-habit faces of sufficiently low energy to provide a stable corner embryo in unsaturated conditions. This contrasts with the case of a liquid surface where all surface orientations have the same energy, so that a liquid corner embryo with concave surface can be stable below the saturation point. It is to be expected that the inclusion of second neighbour interactions might relax this effect.*

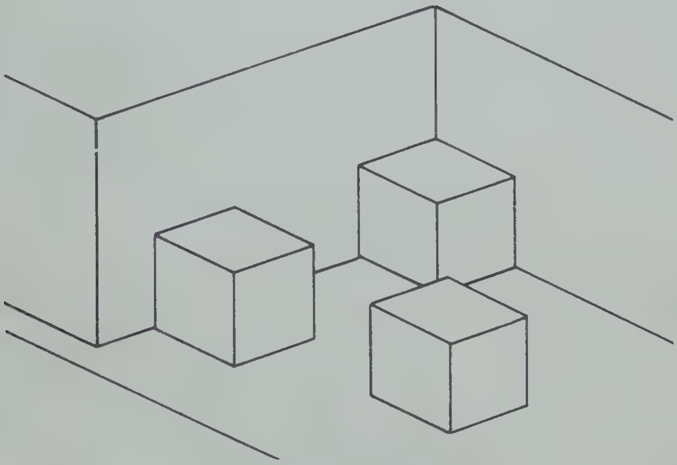


Fig. 4.—Nucleation of a simple cubic embryo on a plane surface, at a step, and in a corner.

Since we are concerned here principally with the case of ice we shall give results similar to equations (**22**)–(**24**) for a hexagonal lattice. The geometry is here rather more complicated, and Wulff's rule is useful in deciding upon the shape of the embryo. The general result is

$$\Delta G^* = \frac{16\sqrt{3}\cdot\sigma^3}{\Delta G_v^2} \cdot f(\gamma), \quad \dots\dots\dots\dots\dots \quad (\mathbf{22})$$

where $f(\gamma)$ depends upon the geometry as shown in Table 1. Again the surface nucleation term has been neglected.

These results are less simple than in the cubic case, but the middle row of the table shows the same sort of behaviour as shown by equation (**21**). In the case of nucleation involving two prism faces meeting at an angle of $\pi/3$ a new phenomenon arises. The factor $f(\gamma)$ becomes zero for $\gamma = \frac{1}{2}$, and for such surfaces no supersaturation is required for nucleation in the edge or corner. If $\gamma < \frac{1}{2}$

---

* The substrate throughout has been considered to have the same habit faces as the embryo. If other faces occur on the substrate and produce a step, then it may become possible for an embryo to be retained on the step under unsaturated conditions, provided that the step angle is sufficiently small and the γ value low enough on the non-habit interfaces.

a change in the sign of ΔG_v is required for the critical embryo to exist, and this implies that a stable embryo can be retained in the edge or corner under unsaturated conditions.

It should perhaps be mentioned that no consideration has been given to the effect of any elastic strain energy caused by coherent growth upon the substrate. The contribution of this effect becomes increasingly serious when nucleation in steps or corners is considered, the strain energy being roughly proportional to the number of contact planes involved. We have, however, little knowledge of the extent to which such coherent growth occurs, and detailed consideration of particular cases would not be fruitful at this time.

TABLE 1

SHAPE FACTOR $f(\gamma)$ FOR VARIOUS SITUATIONS, FOR $\gamma \leqslant 1$

(Included angles are shown for prism faces)

Planes		Steps		Corners	
Free	1	Base-prism	$\gamma^2(2+\gamma)/12$		
Base	$\gamma/2$	Prism-prism $2\pi/3$	$\gamma^2/3$	$2\pi/3$	$\gamma^3/6$
Prism ..	$\gamma(2+\gamma)/6$	Prism-prism $\pi/3$	$(2\gamma-1)^2/6$	$\pi/3$	$\gamma(2\gamma-1)^2/12$

V. CRYSTAL GROWTH

We have as yet considered only the nucleation of a crystal embryo, and in the latter part of the discussion have neglected the necessity for nucleation of new crystal layers upon the growing embryo. We now return to this aspect of the problem and examine the course by which an embryo grows to a macroscopic crystal.

Reverting to the cubic crystal habit for simplicity, we recall that the free energy barrier to nucleation, at the intersection of n substrate planes, is

$$\Delta G^*_{[3]} = \frac{32\sigma^3}{\Delta G_v^2}\left(\frac{\gamma}{2}\right)^n . \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (23)$$

If now we examine the minimum free energy barrier to the formation of a two-dimensional embryo on top of such a three dimensional embryo, we find in a very similar manner

$$\Delta G^*_{[2]} = -\frac{4a\sigma^2}{\Delta G_v}\left(\frac{\gamma}{2}\right)^{n-1} , \qquad \ldots\ldots\ldots\ldots\ldots \quad (24)$$

where a is the lattice spacing in the embryo. (This result holds if $n \geqslant 1$; if $n=0$, we must take $\gamma=2$.)

If then a three-dimensional embryo has nucleated, it will be able to grow freely by two-dimensional nucleation provided

$$\Delta G^*_{[2]} < \Delta G^*_{[3]}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (25)$$

which may be reduced to

$$-\Delta G_v < 4\sigma\gamma/a. \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (26)$$

The parameter γ has so far been considered to be entirely arbitrary within the range $0 \leqslant \gamma \leqslant 2$, but a little consideration shows that γ cannot in fact be made arbitrarily small, since the thickness of the embryo in any dimension must be at least equal to the lattice distance a. Equations (**6**) and (**7**) therefore lead to

$$h = -2\sigma\gamma/\Delta G_v \geqslant a, \qquad \ldots\ldots\ldots\ldots\ (\mathbf{27})$$

or

$$-\Delta G_v \leqslant 2\sigma\gamma/a. \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ (\mathbf{28})$$

If (**28**) is satisfied, as it must be, then (**26**) is automatically satisfied and the embryo can grow by two-dimensional nucleation of new crystal planes.

A growth mechanism such as this, however, becomes inactive when the value of $|\Delta G_v|$ has been decreased by diffusion of vapour to growing crystals. Crystals which contain screw dislocations require much lower supersaturations for appreciable growth to occur and will therefore grow at a much faster rate than undislocated embryos. In the absence of any other source of dislocations, embryos nucleated on substrate steps may grow slowly until they cross a screw dislocation in the substrate. This dislocation structure is then transferred to the growing crystal and further growth by the dislocation mechanism may proceed.

VI. ENTROPY EFFECT

Since we are applying our discussion to the nucleation of ice crystals, mention should be made of an effect suggested by the author (Fletcher 1959) which may have considerable influence on the nucleation efficiency of certain ionic substrates.

The dipoles of ice molecules normally have random orientations, even at absolute zero, so that ice has a residual entropy. If ice is caused to grow on an ionic substrate of which all the exposed ions are of one sign, then the configurational entropy of the ice molecules is greatly reduced and the free energy of the ice embryo increased.

This effect appears to be serious enough in the case of basal faces of substances like silver iodide and lead iodide to lower the threshold for the nucleation of freezing by about 30 °C and effectively render these faces inactive as nucleating agents. On the other hand the prism faces of these materials contain equal numbers of ions of each sign and after the first ice layer has grown the entropy has its normal value. We should thus expect these prism faces to provide the active nucleating sites.

Large steps on basal planes will, of course, expose prism faces, so that this theory, too, predicts that steps and corners on basal faces provide good nucleating sites, though for an entirely different reason.

VII. EXPERIMENTAL RESULTS

Many observational studies have been made on the nucleation of ice crystals upon substrates of silver iodide or lead iodide. Among these should be mentioned the work of Jaffray and Montmory (1956, 1957), of Kleber and Weis (1958), and of Mason (1959) ; similar work has also been done by the present author.

The experiments all agree that steps in surfaces are preferred nucleation sites and that nucleation on perfect surfaces requires very large supersaturations or supercoolings. Mason reported that dislocations appeared to be good nucleation sites, but examination of his photographs shows relatively large pits around the dislocation site so that large steps were probably also available for nucleation.

Plate 1 shows the growth of ice crystals upon a substrate of lead iodide at −15 °C, growth taking place from a supersaturated vapour. It can be seen that steps provide numerous nucleation sites, the ice crystals near the centres of the terraces also appearing to be on steps rather than at the true centre where the dislocation line emerges. It should be noted that the steps in this photograph are of the order of a micron in height, so that they are of sufficient size for our theory to be applicable. Steps only a few molecular layers in height would require a different theoretical treatment.

It does not seem possible to decide on the basis of the experimental evidence whether the calculated entropy effect takes place or not. The results could be explained on the basis of operation of the entropy effect together with a small γ value, or of no entropy effect and a considerably larger value of γ. A decision on this point must await more definitive experiments.

VIII. Conclusions

We have seen how the classical theory of the nucleation of new phases on perfect substrates can be extended to discuss the effects of a variety of crystalline surface imperfections. Dislocations meeting the surface, whether of edge or screw character, do not ordinarily provide preferred nucleation sites, but steps and re-entrant corners are sites in which the free energy barrier to nucleation is greatly reduced.

An effect of this nature is to be expected by analogy with the nucleation of liquid droplets, but some of the features associated with the retention of liquid embryos in concave edges under unsaturated conditions do not apply fully to crystals. Crystals of hexagonal habit may, however, be retained in sufficiently sharp concave steps provided the interfacial free energy is small enough.

There is not yet sufficient evidence to allow a decision to be made on the reality of an entropy effect in ice crystal nucleation previously suggested by the author, but this does provide one of two alternate explanations of observed nucleation behaviour.

IX. References

Becker, R., and Doring, W. (1935).—*Ann. Phys., Lpz.* (5) **24**: 719.

Burton, W. K., and Cabrera, N. (1949).—*Disc. Faraday Soc.* **5**: 33.

Fletcher, N. H. (1958).—*J. Chem. Phys.* **29**: 572.

Fletcher, N. H. (1959).—*J. Chem. Phys.* **30**: 1476.

Frank, F. C. (1949).—*Disc. Faraday Soc.* **5**: 48.

Frank, F. C. (1951).—*Acta Cryst.* **4**: 497.

Giddings, J. C., and Eyring, H. (1958).—*J. Phys. Chem.* **62**: 305.

Jaffray, J., and Montmory, R. (1956).—*C.R. Acad. Sci., Paris* **243**: 126, 891.

Jaffray, J., and Montmory, R. (1957).—*C.R. Acad. Sci., Paris* **244**: 859.

Kleber, W., and Weis, J. (1958).—*Z. Kristallogr.* **110**: 30.

Mason, B. J. (1959).—*Weather* **14**: 81.

Penny, A. H. A. (1948).—*Proc. Camb. Phil. Soc.* **44**: 423.

Turnbull, D. (1950).—*J. Chem. Phys.* **18**: 198.

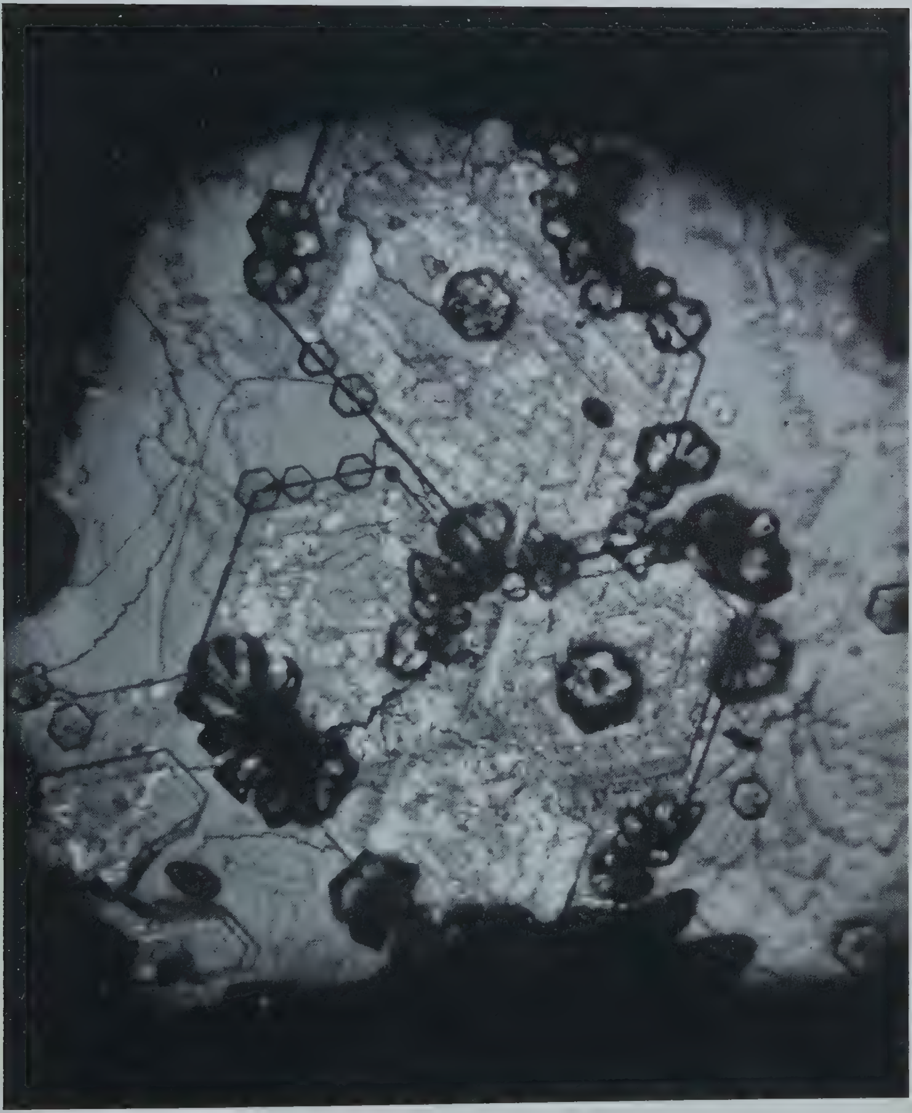
Turnbull, D., and Fisher, J. C. (1949).—*J. Chem. Phys.* **17**: 71.

Turnbull, D., and Vonnegut, B. (1952).—*Industr. Engng. Chem.* **44**: 1292.

Volmer, M., and Weber, A. (1925).—*Z. Phys. Chem.* **119**: 277.

Wulff, G. (1901).—*Z. Kristallogr.* **34**: 449.

NUCLEATION AND GROWTH OF ICE CRYSTALS



Nucleation of ice crystals from the vapour onto a lead iodide crystal substrate at about —15 °C.

THE PHOTOGRAPHIC PROCESS AS A DIATOMIC REACTION

By C. Candler*

[*Manuscript received September* 25, 1959]

Summary

If sensitive areas exist in unexposed crystals, the number of grains which develop on exposure is given by the Poisson law, and this conflicts with experiment; but, if a development centre is formed in a diatomic reaction triggered by P trapped electrons, the fraction n of the grains capable of development after an exposure E is

$$n/(1-n)=(sE)^P,$$

where s is the reaction cross section.

Examination of the papers of earlier workers shows that $\log\{n/(1-n)\}$ is linear in $\log E$ for any one grain size, so theory is confirmed.

The photon number P increases with the size of crystal, but discontinuously. Non-integral values of P cannot be attributed to a mixture of grains each with P integral, but are inherent in the constitution of a single grain.

The reaction cross section s can be accurately calculated, when the exposure is measured in absolute units, for the theory of dimensions requires that E be expressed in photons per unit area. In two emulsions s is proportional to the volume of the crystal.

I. INTRODUCTION

Forty years ago many physicists rightly held that the key to the photographic process lies in the exposure curve of an emulsion spread so thin that absorption plays no part. And although grain counts are laborious, five emulsions (Slade and Higson 1921; Toy 1921, 1922, 1926; Clark 1923, 1924; Sheppard, Trivelli, and Loveland 1925; Webb 1948) were studied. But no one was able to explain the counts, and when Silberstein (1928) showed that either the quantum theory or the Poisson distribution must be rejected, grain counts were abandoned as unprofitable.

Fortunately, the grain counts were recorded with meticulous care. And this paper attempts to show that all fit an equation so simple that it must surely have emerged from an empirical study. The theory which best describes the counts is a diatomic reaction between two silver atoms.

II. OLD GRAIN COUNTS RE-EXAMINED

The probable error in the fraction n of the grains which develop can be reduced only by counting large numbers (Slade and Higson 1921). Sheppard's count of 15,000 grains spread over 32 different areas is the most accurate ever

---

* South Australian Institute of Technology, Adelaide.

undertaken. Slade and Higson usually counted 500 grains in each experiment, and Toy, who used grain counts only to obtain an answer to a specific problem, 100.

In a Seed Graflex emulsion with the crystal area running from 0·1 to 6·9 μ^2, Sheppard recorded his counts by size as well as exposure. For areas between 0·9 and 4·5 μ^2 four or more exposures were recorded, and always log $\{n/(1-n)\}$ is closely linear in log E (Fig. 1).

Similar results were obtained by Sheppard (1925) on a slow emulsion (Fig. 2), by Slade and Higson (1921) on a whole emulsion " containing grains as nearly as possible the same size " (Fig. 3), and by Clark (1924) in a comparison of a plate developed normally with one treated with dilute chromic acid (Fig. 4).

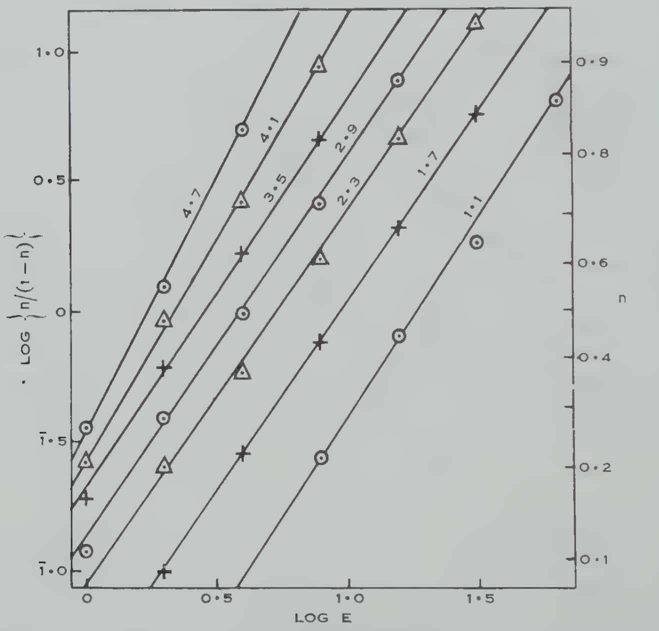


Fig. 1.—The diatomic function log $\{n/(1-n)\}$ for several grain areas in the Seed Graflex emulsion recorded by Sheppard, Trivelli, and Loveland. The grain area in square microns is shown beside each line.

While all these grain counts confirm the linearity of log $\{n/(1-n)\}$, some reveal an unexpected regularity, to wit two straight lines with a break between. Thus, in three successive experiments by Slade and Higson (Fig. 5), both arms are well attested. The angle seems to be quite sharp, and the two lines do not appear to be the asymptotes of a hyperbola. Moreover, the " break " seems to be a common phenomenon. In the smaller grains of the Seed Graflex emulsion (Fig. 6) the graphs may be taken for curves, but bearing in mind the previous figure, two lines with a break between is the more likely interpretation. Again Toy (1922) counted grains (Fig. 7) of four different areas, 0·98, 1·75, 2·73, and 3·93 μ^2, thus achieving a uniformity in size 30 times better than Slade and Higson. Yet a break occurs in two of these lines.

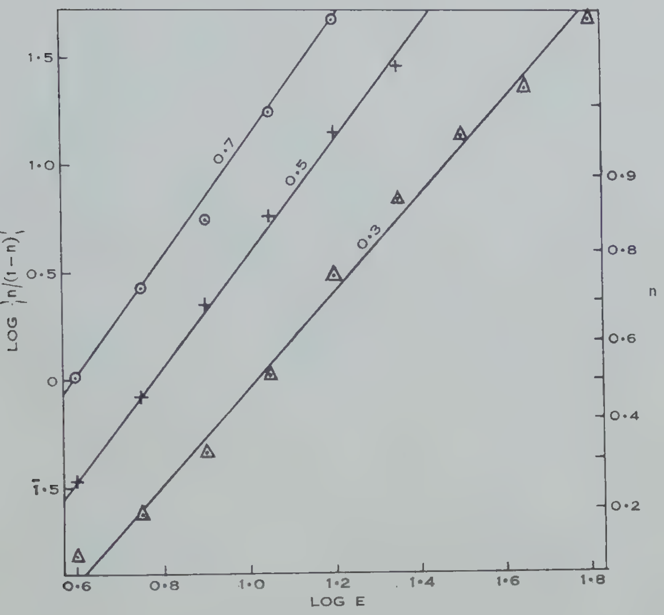


Fig. 2.—The diatomic function in the slow emulsion recorded by Sheppard, Trivelli, and Loveland.

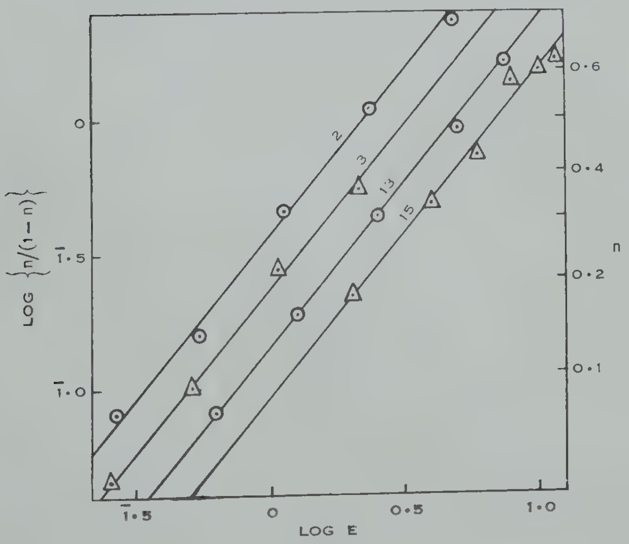


Fig. 3.—Experiments 2, 3, 13, and 15 of Slade and Higson show the diatomic function linear in log *E*. The relative exposure of successive experiments is not at issue, so some of the lines have been displaced laterally.

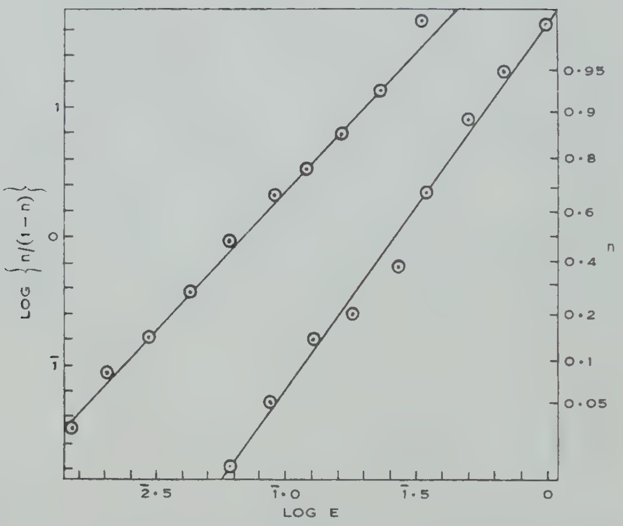


Fig. 4.—An emulsion studied by Clark. The line on the left is after normal development, that on the right after treatment with dilute chromic acid.

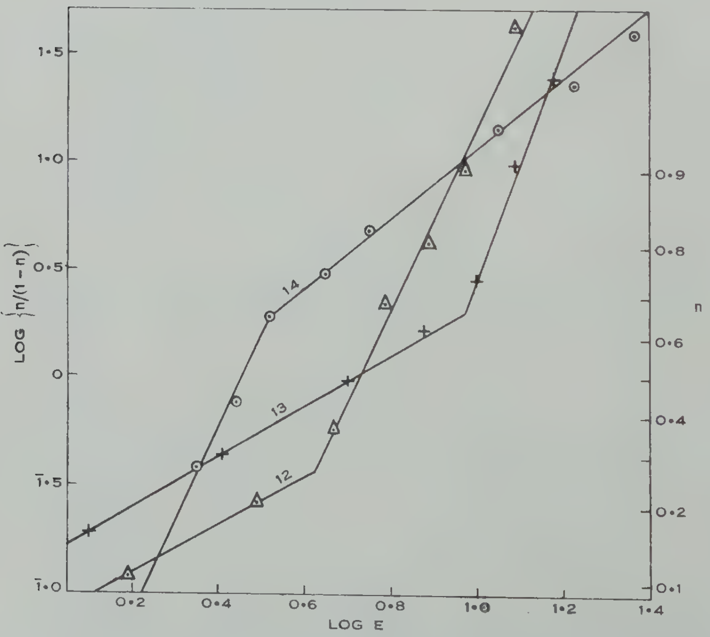


Fig. 5.—Three experiments by Slade and Higson showing well-defined breaks. In these three experiments the intensity remained the same in any one experiment, while the time varied from point to point.

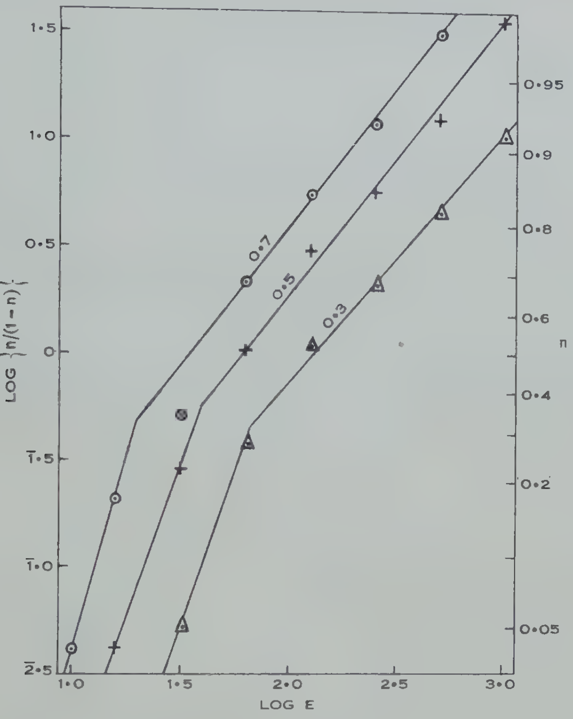


Fig. 6.—The smaller grains of the Seed Graflex emulsion recorded by Sheppard, Trivelli, and Loveland appear to exhibit breaks. Only in the grains of area 0·7 μ^2 is there any indication that the "break" is not sharp.

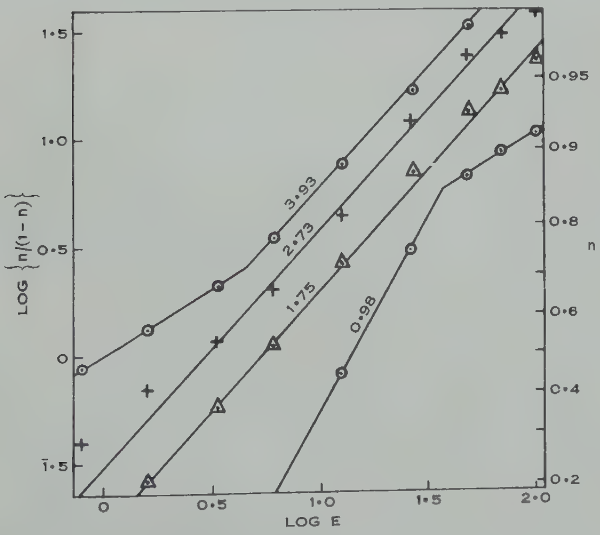


Fig. 7.—The diatomic function for four grain areas selected by Toy. Toy counted fewer grains than Slade and Higson, so the unavoidable statistical error is higher.

III. The Poisson Distribution

From 1920 to 1928 everyone who worked on photographic theory accepted the Poisson distribution as axiomatic, but not one was able to get it to explain his counts.

The reason must be some error in the assumed mechanism. For the Poisson law is a mathematical not a physical law, so the argument is not open to attack once the postulated conditions have been established. And these read : " When m is the average number of pulses striking an object in t seconds, and pulses are independent and random and the chance of a hit rare, the probability that x of the pulses hit the object is $e^{-m}m^{x}/x\,!$ " (Burington and May 1952).

Study of sulphur sensitization led to the belief that the sensitive areas of a crystal are few and small, and only when two or more photons are absorbed in one of these areas is a development centre formed. If this were the true picture, the Poisson law would surely apply. But the converse is equally true ; as the Poisson law does not apply, the picture must be at fault.

As one possible alternative, suppose that no sensitive areas exist before exposure (Baker 1956) ; and instead a development centre is formed when two photons are absorbed contiguous in space and almost coincident in time. There is then no pre-existing " object " and the absorption of the first photon cannot be called a " rare " event, for absorption anywhere in the crystal is effective. Moreover, the two absorptions are not properly mere repetitions of the same event, and on this ground, too, the Poisson law fails.

This argument in no way impugns the validity of the Poisson law in other branches of photographic theory. Whatever the mechanism which produces the first development centre in a crystal, exactly the same mechanism produces the second and third, so the number of specks per grain after partial development satisfies Poisson's law (Svedberg 1921). But an argument from specks per grain to photons per centre lacks any basis in logic.

IV. The Formation of a Development Centre

When radiation falls on a crystal of silver halide, it may form a development centre, and a single centre renders the whole crystal " capable of development ". This is a cumbrous phrase, which irritates when repeated, so here the word " fertilize " will be borrowed from another science, and photons that produce a development centre will be said to fertilize the crystal. One hopes this nomenclature may seem appropriate, for when a sperm fertilizes an ovum, the ovum is " rendered capable of development ".

Thirty years ago Baker (1926) studied the formation of development centres experimentally, and his conclusions still deserve quotation : " Two quanta are concerned, and both must be absorbed within a short interval . . . In the absence of a second absorption, the effect of the first is transitory . . . The two quanta must be taken up by different components of the absorber, for otherwise a single quantum of double the frequency necessary for any action would suffice to cause the whole effect. This would imply a discontinuity in the ultraviolet, and no such discontinuity has been observed. Further the maximum wavelength effective in the two stages must be close to the maximum for any photographic action ". And, in conclusion, " either component may absorb the first quantum, and the probability of either absorbing the first quantum is the same ".

In effect this last sentence states that the two absorbers have the same reaction cross section. And in the last decade nuclear physicists have shown that an absorption cross section identifies a particle as surely as its mass. So to accept the argument is to find oneself forced to enquire whether the two absorbers are not identical.

V. THE MONOLAYER DERIVED FROM A DEEP EMULSION

That the number of silver grains formed in a monolayer can be derived from the characteristic of a deep emulsion has been understood for half a century (Channon 1906 ; Buisson and Fabry 1924). One merely differentiates the number of grains in the coating with respect to the depth. But this approach was of little use until an equation of the characteristic was available.

When the exposure equation of an emulsion so deep that no light is lost through the back surface,

$$A+P\log E=PKD+\log D, \qquad \ldots\ldots\ldots\ldots (1)$$

had been derived from theory (Candler 1959), and had been proved valid in 17 common emulsions, however, differentiation offered a new hope of success.

In this exposure equation P is the " photon number ", a quantity which measures the number of trapped electrons required to fertilize a crystal, while K is the ratio of the absorption cross sections of a halide crystal before and after development,

$$K=a_c/a_g, \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

and by a fortunate chance the values of P and K are unchanged when natural densities and natural logarithms replace common densities and common logarithms.

The number of grains **N** in a cylinder of unit area cut perpendicular to the surface of an emulsion is a measure of the density (Nutting 1913, Trivelli 1946),

$$\mathbf{D}=\mathbf{N}a_g. \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Here **D** is the natural density, defined as ln $(1/T)$, T being the transmittance.

To obtain the superficial grain concentration N in a thin layer, the number of grains in the cylinder is differentiated with respect to the depth z,

$$N=\mathrm{d}\mathbf{N}/\mathrm{d}z=(\mathrm{d}\mathbf{D}/\mathrm{d}z)/a_g. \qquad \ldots\ldots\ldots\ldots\ldots (4)$$

Here the " depth " z is conveniently measured up from some point deep in the emulsion where the density is negligible, rather than down from some arbitrary surface, for then D, E, **N**, and z all increase together. The origin of the z-coordinate need not be precisely defined, as only differences in depth are of any interest.

Both **D** and z are more conveniently related to the exposure than to one another, so we write

$$N=(\mathrm{d}\mathbf{D}/\mathrm{d}E)(\mathrm{d}E/\mathrm{d}z)/a_g, \qquad \ldots\ldots\ldots\ldots (5)$$

and the first differential is then obtained from the exposure equation (2)

$$\mathrm{d}\mathbf{D}/\mathrm{d}E=P\mathbf{D}/E(1+PK\mathbf{D}), \qquad \ldots\ldots\ldots\ldots\ldots (6)$$

while the second derives from the exponential decrease of exposure with depth

$$E=E_0 \exp (N_0 a_c z), \quad \text{(7)}$$

where N_0 is the number of halide crystals per unit volume. Thus,

$$\mathrm{d}E/\mathrm{d}z=N_0 a_c E. \quad \text{(8)}$$

On inserting these results in (5), the fraction n of the grains capable of development in a thin slice at depth z is

$$n=N/N_0=PK\mathbf{D}/(1+PK\mathbf{D}), \quad \text{(9)}$$

or

$$n/(1-n)=PK\mathbf{D}. \quad \text{(10)}$$

Fresh grains are formed chiefly deep in the emulsion, where the density is very low, so, if the exposure equation of the deep emulsion is written

$$E^P \propto \mathbf{D}e^{PK\mathbf{D}}, \quad \text{(11)}$$

the important term is **D**, since the second term differs little from unity. Thus the exposure equation of a monolayer may be written

$$(sE)^P=n/(1-n), \quad \text{(12)}$$

where s is the reaction cross section and, if the law of dimensions is to be satisfied, E must be measured in photons per unit area.

VI. A DIATOMIC REACTION

The ratio $n/(1-n)$ suggests a bi-molecular reaction, so the equation is consistent with a reaction between two halide crystals. But in an emulsion the individual crystals are isolated in the gelatine, so this interpretation makes no physical sense.

However, if $(1-n)$ is the fraction of the crystals still virgin, it is also the fraction of the atoms still capable of fertilizing a fresh crystal. For a second centre, formed in a crystal which already contains one centre, does not increase the number of grains that develop.

So, if fertilization is a diatomic reaction triggered by P trapped electrons, the differential equation is

$$\mathrm{d}n/\mathrm{d}E=k(1-n)^2E^{P-1}, \quad \text{(13)}$$

and on integration this takes the form, previously derived from the deep emulsion,

$$(sE)^P=n/(1-n). \quad \text{(14)}$$

VII. THE INFLEXION OF THE EXPOSURE CURVE

Almost all experiments have relied on relative exposures, for absolute measurement requires apparatus seldom available even in an optical laboratory. So the only point on the exposure curve which can be specified is the point of inflexion, where $\mathrm{d}^2n/\mathrm{d}E^2$ is zero. And this is the point at which the divergence between the experimental and Poisson values of n excited Silberstein's (1928) comment.

Two types of reaction, diatomic and Poisson, will be examined before the predictions are compared, but the mathematics of double differentiation need not be reproduced.

A diatomic reaction satisfying equation (**14**) inflects when

$$n_d=(P-1)/2P. \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{15})$$

The Poisson law is more cumbersome mathematically, because the relationship between n and E cannot be made explicit in P, except in the form of a series, in which P determines the number of terms (Mees 1954, at p. 183). But P may be given successive integral values, and each in turn submitted to double differentiation. When $P=2$, the exposure equation is

$$n_p=1-e^{-sE}(1+sE), \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{16})$$

and inflexion occurs when

$$n_p=(e-2)/e=0\cdot264. \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{17})$$

While, when $P=3$, the exposure equation becomes

$$n_p=1-e^{-sE}(1+sE+\tfrac{1}{2}s^2E^2), \qquad \ldots\ldots\ldots\ldots \quad (\mathbf{18})$$

and inflexion occurs when

$$n_p=(e^2-5)/e^2=0\cdot322. \qquad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{19})$$

The percentage of crystals fertilized at inflexion in these two reactions are collated below.

P	1·5	2	3	4
Poisson ..	—	26·4	32·2	29·5
Diatomic	16·7	25·0	33·3	37·5

In a classic paper Silberstein (1928) criticized the Poisson equation because the empirical percentage of crystals fertilized at the point of inflexion is less than that dictated by theory. He does not, indeed, give the figure on which he relied, but he was using Sheppard's grain counts ; if he examined the exposure curves of grains of medium size, where $P=1\cdot44$, the fraction of the grains developed at inflexion is 15·3 per cent. This figure is so much below the 26·4 per cent. of the Poisson equation that Silberstein was well justified in concluding that the quantum theory and Poisson distribution lead to a result incompatible with the facts.

VIII. The Photon Number—Variation with Grain Size

Only in three emulsions, one slow and two fast, have grains of different size been separately recorded.

In both of Sheppard's emulsions the photon number increases with the size of grain. In the fast Seed Graflex (Fig. 8) the increase was discontinuous, P remaining constant while the area increased from 0·9 to 3·5 μ^2 ; and there is some evidence of other steps at 1·20, 1·66, and 2·00. However, but for the indisputable evidence at 1·44, the readings are not accurate enough to disprove a steady rise elsewhere.

Toy (1922) also used a fast emulsion (Fig. 9), but counted only grains of four selected sizes. Nevertheless, his evidence supports a photon number independent of size over a wide range, and of non-integral values.

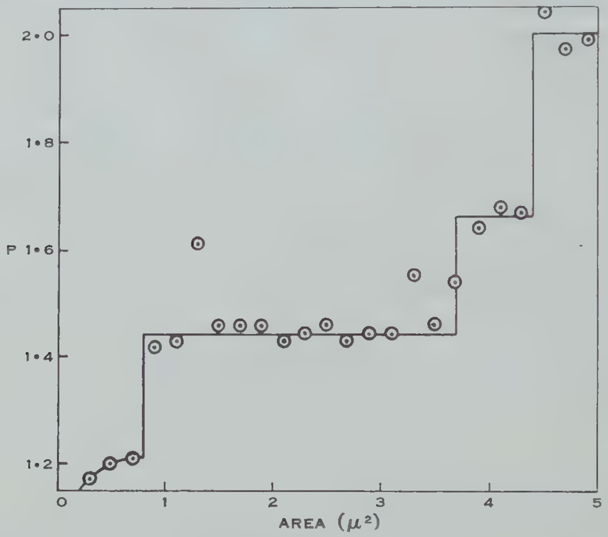


Fig. 8.—Variation of P with the crystal area in the larger crystals of Sheppard's Seed Graflex emulsion.

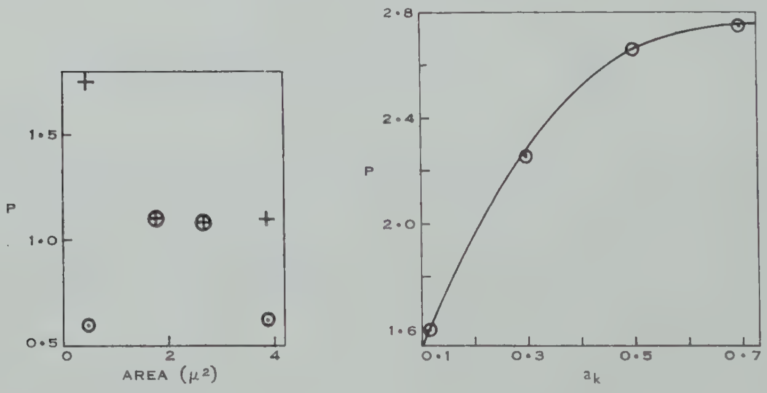


Fig. 9 Fig. 10

Fig. 9.—Variation of P with crystal area in Toy's emulsion. A circle denotes the value at low exposure, and a cross that at high exposure.

Fig. 10.—Variation of the photon number with grain area in Sheppard's slow emulsion.

In Sheppard's slow emulsion (Fig. 10) the rise of P with grain size appears to be continuous, but the percentage increase in size at each step is much larger than in a fast emulsion, so any discontinuity might well appear slurred.

Two other observations suggest that the increase of photon number with grain size is common in other emulsions. Any form of development which aborts a centre will make the small grains incapable of development before the large, for a grain will remain capable of development until the last centre is aborted. So, if the photon number increases with grain size, abortion will increase the effective value of P.

And this is just what occurs when a plate is treated with dilute chromic acid (Fig. 4) (Clark 1924) and when excess bromide is added to the developer (Fig. 11) (Baker 1928). The latter observation was made on a Schumann plate, but in a Schumann plate as in a monolayer absorption plays no part and the fraction of the grains fertilized is D/D_m, where D_m is the maximum density reached on long exposure, so the exposure equation is

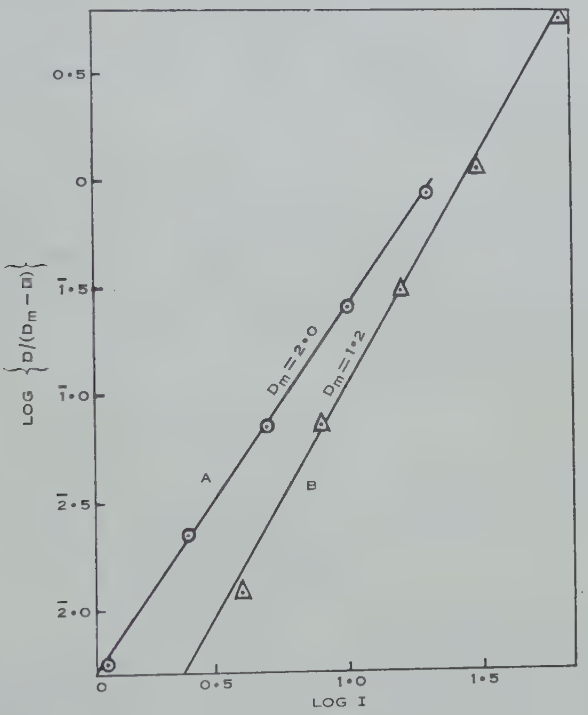


Fig. 11.—Two Schumann plates of the type known as Klein-Hilger with Baker's measurements at 365 mμ. Plate A received full development in metol quinol, and plate B "partial development", bromide being added. The values of P are 1·8 and 2·2 in A and B respectively. The intensity scale is not the same for the two plates.

$$(sE)^P = D/(D_m - D). \qquad (20)$$

This equation accounts very well for Baker's measurements.

IX. The Photon Number Non-integral in a Single Grain

The Poisson law, like the old quantum theory, will only tolerate integral values of P, and for a quarter of a century theorists (Webb 1941; Mitchell and Mott 1957) have tried to explain the characteristic in terms of a mixture of crystals, in each of which P is integral.

That P is non-integral and independent of crystal area from $0 \cdot 9$ to $3 \cdot 5\ \mu^2$ in the Seed Graflex emulsion is some evidence that the value of $1 \cdot 44$ is not due to mixed grains with $P=1$ in some and $P=2$ in others. But on so important an issue, a more rigid proof is desirable.

So consider a monolayer comprised of crystals, characterized by sensitive areas s', s'', and by photon numbers P', P''. If the crystals are present in the proportions f', f'', where

$$f'+f''=1, \qquad (21)$$

and fractions n', n'' are fertilized by an exposure E, the fraction of the grains found developed in the composite emulsion is

$$n=f'n'+f''n'', \qquad (22)$$

for under the microscope the two types of grain are indistinguishable.

The two types of crystal are supposed not to influence each other in any way, being separated by gelatine and spread in a layer so thin that absorption plays no part; and the fractions of the grains fertilized are therefore those given by the normal diatomic equation:

$$n'=(s'E)^{P'}/\{1+(s'E)^{P'}\}, \qquad (23)$$

$$n''=(s''E)^{P''}/\{1+(s''E)^{P''}\}. \qquad (24)$$

On combining these equations,

$$\frac{n}{1-n}=\frac{f'(s'E)^{P'}+f''(s''E)^{P''}+(s'E)^{P'}(s''E)^{P''}}{1+f''(s'E)^{P'}+f'(s''E)^{P''}}. \qquad (25)$$

The grains of medium size in the Seed Graflex emulsion provide a critical test, for with the empirical photon number $1 \cdot 44$ the photon number of the grains may be assumed to be either $P'=1$ or $P''=2$, while f' and f'' cannot differ much from $\frac{1}{2}$. Indeed, as $\log\{n/(1-n)\}$ is linear when $P=1 \cdot 44$, one may reasonably assume it is also linear when $f'=f''=\frac{1}{2}$, and use this to simplify the equation.

As the grains of a count were all the same size, and the reaction cross section is proportional to the grain volume, it is probably the same in all crystals, and we may write $s'=s''=s$. The exposure equation then takes the simple form

$$\frac{n}{1-n}=sE\,\frac{1+sE+2s^2E^2}{2+sE+s^2E^2}. \qquad (26)$$

The curve of $\log\{n/(1-n)\}$ against $\log sE$ is S-shaped (Fig. 12), with the slope asymptoting to unity at low and high exposures, and having the value $1 \cdot 5$ at the mid point where $n=\frac{1}{2}$. If the empirical curves are taken as linear from $n=0 \cdot 02$ to $0 \cdot 98$, this curve can be definitely excluded. But in grain counts the probable error is rather large when n is less than $0 \cdot 05$ or greater than $0 \cdot 95$, so one should probably conclude that, while a photon number integral in each grain is improbable, the possibility has not yet been rigorously excluded.

In some of Slade and Higson's experiments the linearity of log $\{n/(1-n)\}$ is also clear and, as the empirical photon number is 1·20, the argument is only slightly less sure: and, if a non-integral value is accepted in two emulsions, there is scant reason for retaining it elsewhere.

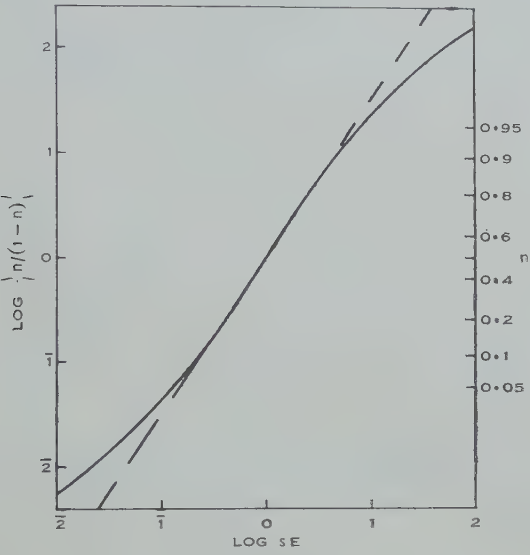


Fig. 12.—The diatomic function in an emulsion whose grains have $P=1$ and $P=2$ in equal numbers. An empirical curve from the Seed Graflex emulsion is drawn broken.

X. The Reaction Cross Section

Before the exposure equations of deep emulsion and monolayer were known, the sensitive area had been estimated as 10^{-3} to 10^{-4} of the area of a crystal (Silberstein and Trivelli 1945). But precise measurement in a given emulsion had never been achieved, so the variation of sensitive area with crystal size has never been studied.

None the less many have asked whether the sensitive specks are distributed through the body of the crystal, or confined to the surface. If the interior of a crystal is as sensitive as the surface, the reaction cross section should be proportional to $a_k^{3/2}$, where a_k is the projected area of the crystal: so a graph of log s against log a_k should have a slope of 1·5, while, if only the surface is sensitive, the slope should be unity.

In two fast emulsions (Figs. 13, 14) the slope is close to 1·5, so the interior is as sensitive as the surface. In Sheppard's slow emulsion, however, the slope differs little from unity (Fig. 15), but the crystals differed so little in size that this result can hardly be considered reliable.

The sensitive areas have usually been supposed to lie all on the surface, since Svedberg (1921) showed that the number of specks of silver formed per grain on partial development is proportional to the area. But the experiments themselves are not in conflict, when properly interpreted. For a development centre consists of two atoms, and a speck visible in the electron microscope of several hundred. So, if the centres near the surface of the grain grow much more rapidly than centres not so readily accessible to the developer, the two results are in good accord.

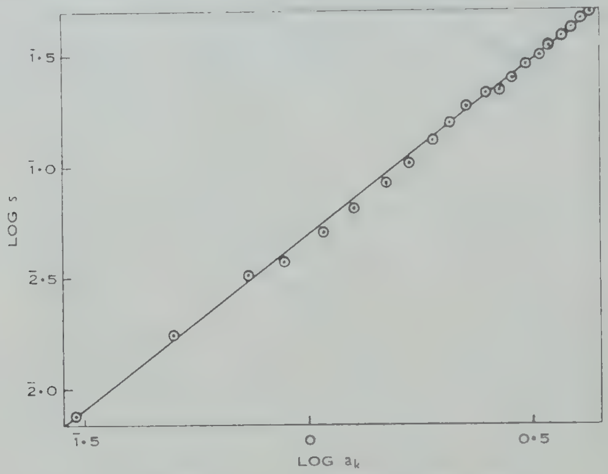


Fig. 13.—Variation of the reaction cross section with crystal area in Sheppard's Seed Graflex emulsion. The slope of the line is 1·5.

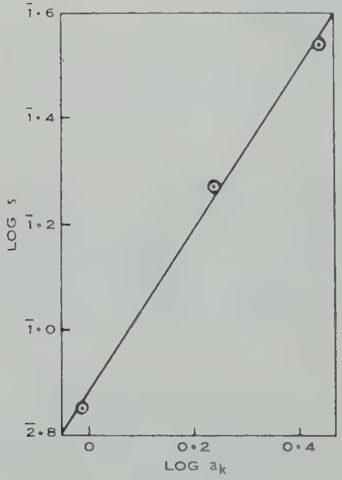


Fig. 14

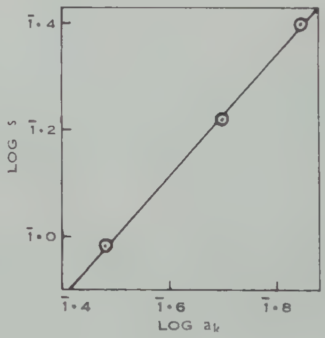


Fig. 15

Fig. 14.—Variation of the reaction cross section with crystal area in a fast emulsion studied by Toy. The slope of the line is 1·5. A fourth point here omitted would make the slope greater than 1·5.

Fig. 15.—Variation of the reaction cross section area in Sheppard's slow emulsion. The slope of the line is 1·13.

Absolute measurements of the reaction cross section are equally interesting. In the exposure equation $n/(1-n)$ is dimensionless, so the product sE must also be a numeric, and this is scarcely possible unless the exposure is measured in photons per unit area. This decided, the reaction cross section can be measured with an error little greater than that involved in the measurement of an exposure in ergs/cm². And, if $E_{\frac{1}{2}}$ is the exposure in ergs/cm² required to fertilize half the crystals,

$$s = 1/5\cdot086\lambda E_{\frac{1}{2}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (27)$$

where λ is the wavelength in millimicrons, and s is the reaction cross section in square microns.

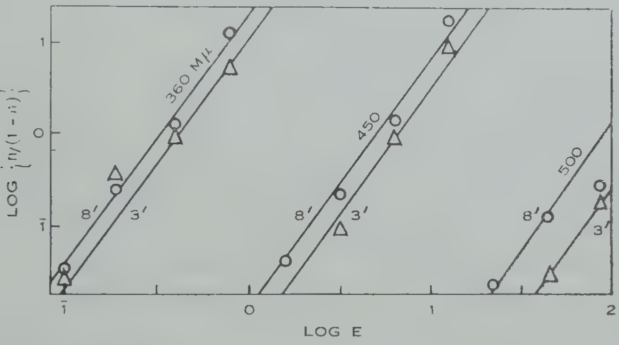


Fig. 16.—A monolayer recorded by Webb in 1948. The slope P is 2·8 at all three wavelengths, whether the development is continued for 3 or 8 min. The exposure E is in ergs/cm².

Only one monolayer (Webb 1948) has had its grains counted after a known absolute exposure (Fig. 16), but here the reaction cross section was in the range found by Silberstein and Trivelli (1945) for the sensitive area (Table 1).

TABLE 1

THE REACTION CROSS SECTION IN SQUARE MICRONS IN THE EMULSION RECORDED BY WEBB

Wavelength (mμ)	Development Time	
	3 min	8 min
360	$1\cdot3\ 10^{-3}$	$1\cdot8\ 10^{-3}$
450	$6\cdot5\ 10^{-5}$	$8\cdot1\ 10^{-5}$
500	$3\cdot6\ 10^{-6}$	$6\cdot0\ 10^{-6}$

XI. THE MECHANISM OF FERTILIZATION

According to Hamilton, Hamm, and Brady (1956) a developed grain results from a reaction which occurs in four steps:

(1) an electron is liberated from a bromide ion into the conduction band,
(2) the electron is trapped,
(3) the electron is neutralized by an interstitial silver ion to form a silver atom,
(4) chemical development is initiated by the speck of silver thus formed.

The first two steps are standard Gurney-Mott (Gurney and Mot 1938, Mott and Gurney 1940) theory and are everywhere accepted, but the third and fourth are still live issues (Mitchell and Mott 1957).

The traps are probably interstitial silver ions, so that steps (2) and (3) are not to be distinguished; and, following Katz (1949) the interstitial silver atom thus formed may be identified as the sub-latent image (Webb and Evans 1940), which has been widely accepted as offering the most satisfactory explanation of reciprocity failure. The sensitizing action of silver sulphide is then likely to be due to strains produced in the lattice, strains which allow the silver ion to move more easily into an interstitial position.

Hamilton's stage (4) accounts well enough for the experiments on carrier motion, which he describes, but it needs modification if it is to explain reciprocity failure, for the latter, being due to the time interval between the arrival of two photons (Webb and Evans 1940, Maerker 1954), is clear evidence that two trapped electrons are involved. However, if one modifies Hamilton's account to read that two trapped electrons react with two silver atoms, and the latter constitute the focus which initiates chemical development, the reason why the two photons must be contiguous in space and nearly coincident in time is evident.

This account helps us to understand the all-or-none nature of development. In commercial developers " a grain is either developed completely or not at all ", wrote Berg (1942). While a theory postulating two or three stages in image formation (Mitchell and Mott 1957) may be compatible with this fundamental fact, it hardly places the all-or-none rule in the centre of the stage, as does Hamilton's picture.

XII. The Significance of the Photon Number

The final problem is the significance of the photon number. Reciprocity failure at low intensity shows that two electrons are needed to fertilize a crystal, but there is no evidence of a larger number. And similarly the diatomic reaction is one between two atoms, not between three or four. Yet empirically the photon number may have any value between 1·0 and 3·5, lying between 1·8 and 2·5 in most emulsions.

The reason is not understood, but a chemical analogy may make the conflict easier to accept. The " order " of a chemical reaction (Hinshelwood 1940) often differs from the number of molecules reacting and may be non-integral. So in an emulsion one may reasonably assume that a development centre is always formed in a reaction between two trapped electrons and two interstitial silver ions, and that the photon number corresponds to the order of a chemical reaction. Formally this is the simplest account, and economy of hypothesis has much to recommend it, until such time as experiment shows just where simplicity fails.

Note added in proof

Thanks to the kindness of Kodak Australasia Ltd. I have recently been able to examine a French periodical not available in Adelaide, and two further points emerge.

The equation here given for the Schumann plate also describes the X-ray characteristic of the B2 emulsion recorded by Tellez-Plasencia (1954). With his four other emulsions a plot of $\log D/(D_s - D)$ is curved, if one accepts the author's " Q " as a true measure of $(1-n)$, but, if in A1 the fog was really 0·04, instead of the 0·02 recorded, and in the emulsions K, C3, and D4 the saturation densities were slightly greater than those recorded, all the plots are linear.

Further, if the equation of the monolayer is integrated, the Hurter equation of the deep emulsion emerges

$$(sE)^P = e^{PK\mathbf{D}} - 1.$$

This is indistinguishable from the author's (Candler 1959)

$$(sE)^P = PK\mathbf{D}e^{PK\mathbf{D}}$$

at low densities, but at all critical points the advantage is with the Hurter equation. I call this the " Hurter equation " because it can be derived equally well from the differential equation given by Hurter (Hurter and Driffield 1890) 70 years ago, if one small correction is permitted.

The two equations of monolayer and deep emulsion appear to describe not only the whole electromagnetic spectrum from 8100 to 0·25 Å, but also the high speed electrons, whose characteristics were recorded by Colombié (1959).

XIII. Acknowledgments

Substantially the whole of this paper was written in Baghdad, and it is a pleasure to acknowledge the encouragement I received from Dr. Sheeth Nu'mann. For 20 years before the revolution of July 1958, Dr. Nu'mann had been Director General of Industry in Iraq.

XIV. References

Baker, E. A. (1926).—*Proc. Roy. Soc. Edinb.* **47**: 34.
Baker, E. A. (1928).—*Proc. Roy. Soc. Edinb.* **48**: 124.
Baker, E. A. (1956).—*J. Photogr. Sci.* **5**: 94.
Berg, W. F. (1942).—*Rep. Progr. Chem.* **39**: 49.
Buisson, H., and Fabry, C. (1924).—*Rev. Opt. (Théor. Instrum.)* **3**: 1.
Burington, R. S., and May, D. C. (1952).—" Handbook of Probability and Statistics." (Handbook Pub.: Sandusky, Ohio.)
Candler, C. (1959).—*Appl. Spectros.* **13**: 97.
Channon, H. J. (1906).—*Photogr. J.* **46**: 216.
Clark, W. (1923).—*Photogr. J.* **63**: 230.
Clark, W. (1924).—*Photogr. J.* **64**: 91.
Colombié, N. (1959).—*Sci. Ind. Phot.* **30**: 165.
Gurney, R. W., and Mott, N. F. (1938).—*Proc. Roy. Soc.* A **164**: 151.
Hamilton, W. J., Hamm, F. A., and Brady, L. E. (1956).—*J. Appl. Phys.* **27**: 874.
Hinshelwood, Sir Cyril (1940).—" The Kinetics of Chemical Change." p. 71 f. (Oxford Univ. Press.)
Hurter, F., and Driffield, V. C. (1890).—*J. Soc. Chem. Ind.* **9**: 455. (Memorial Volume 1920: 109.)
Katz, E. (1949).—*J. Chem. Phys.* **17**: 1132.
Maerker, R. E. (1954).—*J. Opt. Soc. Amer.* **44**: 8.
Mees, C. E. K. (1954).—" Theory of the Photographic Process." p. 183. (Macmillan: New York.)
Mitchell, J. W., and Mott, N. F. (1957).—*Phil. Mag.* **2**: 1149.

MOTT, N. F., and GURNEY, R. W. (1940).—"Electronic Processes in Ionic Crystals." (Oxford Univ. Press.)

NUTTING, P. G. (1913).—*Phil. Mag.* **26** : 43.

SHEPPARD, S. E., TRIVELLI, A. P. H., and LOVELAND, R. P. (1925).—*J. Franklin Inst.* **200** : 51.

SILBERSTEIN, L. (1928).—*Phil. Mag.* **5** : 464.

SILBERSTEIN, L., and TRIVELLI, A. P. H. (1945).—*J. Opt. Soc. Amer.* **35** : 93.

SLADE, R. E., and HIGSON, G. I. (1921).—*Proc. Roy. Soc.* A **98** : 154.

SVEDBERG, T. (1921).—*Photogr. J.* **62** : 186.

SVEDBERG, T. (1922).—*Photogr. J.* **62** : 310.

TELLEZ-PLASENCIA, H. (1954).—*Sci. Ind. Phot.* **25** : 425.

TOY, F. C. (1921).—*Proc. Roy. Soc.* A **100** : 111.

TOY, F. C. (1922).—*Phil. Mag.* **44** : 352.

TOY, F. C. (1926).—*Brit. J. Photogr.* **73** : 704.

TRIVELLI, A. P. H. (1946).—*J. Franklin Inst.* **240** : 1.

WEBB, J. H. (1941).—*J. Opt. Soc. Amer.* **31** : 559.

WEBB, J. H. (1948).—*J. Opt. Soc. Amer.* **38** : 312.

WEBB, J. H., and EVANS, C. H. (1940).—*Photogr. J.* **80** : 188.

ZONE THEORY OF LIQUIDS

By R. E. B. Makinson* and A. P. Roberts*

[*Manuscript received September 25, 1959*]

Summary

A one-dimensional model is considered in which the potential energy of an electron is nearly periodic over short distances but has no long-range order.

The number of states of the electron per unit energy range has been calculated numerically for a number of chains of identical delta-function potential wells, the number of wells being usually 2000. The distance between adjacent wells had a cut-off parabolic distribution about the mean value, various values being taken for the r.m.s. deviation σ.

It was found that (for positive energies) there is a definite energy gap in which no states appear provided σ is small enough.

To throw light on the question of how much of the gap narrowing is produced by the short-range disorder inherent in the model and how much by the lack of order over long ranges, some of the computations were repeated with the short-range disorder smoothed in increasing degree. The narrowing of the gap was thereby progressively reduced. It was found also that similar computations with a model having long-range order but with variation of the position of the wells about ordered positions exhibited narrowing of the gap comparable to that mentioned above.

The relation of these results to the conclusions of Gubanov and of Landauer and Helland is briefly discussed.

I. Introduction

The quantum mechanical treatment of the motion of an electron in a perfectly periodic crystalline field shows that the energy may fall only in allowed bands and leads to a successful explanation of many electrical and thermal conduction properties of metals, semiconductors, and insulators. The finding that melting does not in general greatly affect these properties suggests that a similar theory of zones should be deducible for lattices possessing short-range order only, as in liquids and glassy solids.

Towards building up such a theory several authors have considered one-dimensional models (Gubanov 1954; Landauer and Helland 1954; Eisenschitz and Dean 1957).

Gubanov's model is a potential $V(x)$ in which a series of equal wells have mean spacing a but the distance between each and the next is $a(1+\varepsilon\gamma)$, $\varepsilon \ll 1$ and γ has a Gaussian distribution. Long-range order is thus lost and there is additionally a small disturbance of short-range order. By choosing a distorted coordinate scale ξ in which $V(\xi)$ becomes periodic and solving approximately the equation resulting from Schrödinger's, with ε as a parameter of smallness, Gubanov concludes that the zone structure for $\varepsilon=0$ is preserved but that all levels are raised proportionately to ε^2. The lower boundary of each allowed zone is found to rise by less than the upper boundary, so that the forbidden zones contract.

---

* Physics Department, University of Sydney.

In these calculations, however, there is a lack of rigour in making the necessary approximations which makes alternative methods of studying such a model desirable. Furthermore, no criterion is available from Gubanov's treatment for the amount of disorder necessary to make the energy gap disappear.

One finds that straightforward perturbation of the potential from zero by a small function $\lambda V(x)$, with $V(x)$ as in Gubanov's model, gives interesting results (which will be discussed elsewhere) which are not, however, applicable near the edge of a band where the main interest lies. Reliable extension to the band edge by methods similar to those used for a periodic lattice (e.g. Mott and Jones 1936, p. 59 ; Kittel 1956, p. 583) has not yet been achieved.

Eisenschitz and Dean (1957) have derived electronic energy levels in a tight-binding approximation for a disordered chain. However, they found it necessary to assume that because off-diagonal elements in the energy matrix are on the average small they may be neglected, in which case perturbation treatment of the tight-binding approximation is in any case adequate. This neglect, however, lacks proper justification and there is a missing tail of unknown length on the energy distribution curves.

Discussion of the eigenvalues of Schrödinger's equation with $V(x)$ as above seems to be a problem beyond the present stage of rigorous development of stochastic theory.

Interest therefore attaches to purely numerical calculations for long chains of potential wells exemplifying in various ways lack of long-range order.

Landauer and Helland (1954) have done such calculations for two short chains (50 " atoms ", each a square well), relating the number of stationary states per unit energy range to the energy. This seems, however, an inadequate basis for general conclusions.

An important question for which a reliable answer is required is whether in an infinite chain, lacking long-range order, there may be energy gaps in which there are no states whatever, or whether the states are always thinly spread with a density approaching zero as the range of order increases to infinity.

We have accordingly taken a model like that of Kronig and Penney (1931) of equal delta-function potential wells with uncorrelated spacing distributed about a mean. The length of the chain, usually 2000 " atoms ", was always sufficient to ensure that long-range order was breached many times over within its length, and that additional length brought no further qualitative changes in the results.

II. Numerical Computations

(a) *Long- and Short-range Disorder together*

A chain of δ-function potentials, each of strength $-2\varkappa_0\hbar^2/(2m)$, situated at $x_0=0$, x_1, . . ., $x_{N-1}=L$, gives the one-electron Schrödinger equation

$$\psi''+[\varkappa_0^2\nu^2+2\varkappa_0\Sigma\delta(x-x_n)]\psi=0,$$

where the electron energy $E=\varkappa_0^2\nu^2\hbar^2/(2m)$.

We wish to find $N(\nu)$, the number of eigenstates per unit length with an energy less than $\varkappa_0^2\nu^2\hbar^2/(2m)$, i.e. the integrated density spectrum. It is a well-known result (e.g. Rojansky 1938, p. 177) that with suitable boundary conditions $LN(\nu)$ is simply the number of nodes in $\psi(x, \nu)$, and hence we perform a node count.

With the definitions

$$a_n = \lim_{x \to x_n - 0} \left(\frac{\psi'}{\varkappa_0 \nu \psi} \right) = -\tan \chi_n,$$

$$b_n = \lim_{x \to x_n - 0} \left(\frac{\psi'}{\varkappa_0 \nu \psi} \right) = -\tan \varphi_n,$$

with

$$-\tfrac{1}{2}\pi \leqslant \varphi_n \leqslant \tfrac{1}{2}\pi\,; \quad u_n = (x_{n+1} - x_n)/a, \;\; a = \langle x_{n+1} - x_n \rangle, \;\; \lambda = \varkappa_0 a,$$

the connection requirements on ψ and ψ' yield

$$b_n = a_n - 2/\nu,$$
$$\chi_{n+1} = \varphi_n + \nu\lambda u_n.$$

It follows, as Lax and Phillips (1958) (with some slight differences in notation), that the number of nodes in the nth cell is simply* the integral part of $(\chi_n + \frac{1}{2}\pi)/\pi$.

It is apparent from the above that the spectrum is determined (in a long chain) by the dimensionless quantities ν^2 (an " energy "), λ (a " potential strength "), and u (a stochastic " lattice spacing ").

To achieve the destruction of long-range order, successive u were selected independently and at random. Cut-off parabolic distributions were chosen :

$$\mathrm{d}P(u) = \frac{3\sqrt{5}}{20\sigma} \left[1 - \frac{(u-1)^2}{5\sigma^2} \right] \mathrm{d}u,$$
$$1 - \sigma\sqrt{5} \leqslant u \leqslant 1 + \sigma\sqrt{5}.$$

Thus $\langle u \rangle = 1$, and the standard deviation of u is σ.

In such a chain, the r.m.s. value of the deviation of x_n from its expectation value na, expressed as a fraction of a, is $\sigma n^{\frac{1}{2}}$. Hence long-range order may be considered to be essentially destroyed in about $(4\sigma^2)^{-1}$ cells.

For each σ, a node-count for various energies was performed as described above, on the digital computer SILLIAC. A random number routine generated variate values rectangular in $(-1, 1)$, from which a variate u having the appropriate parabolic distribution was simply obtained.

The same set of 2000 successive u was used for all energies, and thus the results for a given σ refer to the one sample chain. About two minutes of machine time was needed to give $N(\nu)$ for one value of ν.

The initial boundary condition $a_0 = 1$ was adhered to throughout. No final boundary condition was imposed ; the consequence of this lack is an arbitrariness of ±1 in the node count. (If ψ_1, ψ_2 are solutions satisfying different boundary conditions, the constancy of the Wronskian implies that their zeros must alternate with each other.)

---

*This is simpler than the statement by Lax and Phillips (1958) because of our stipulation $-\frac{1}{2}\pi \leqslant \varphi_n \leqslant \frac{1}{2}\pi$.

Counts were printed out for the first 1000 cells for comparison with the results for 2000 ; this disclosed that the sampling fluctuation is of the same order as the node count error—about 0·001 in $N(\nu)$. As a further check against the possibility that $N(\nu)$ might vary systematically as the length of the chain increased (corresponding to the more pronounced violation of long-range order), a few runs were made with 4000 and 8000 cells ; the results were found to agree with those for 2000 within the error mentioned above.

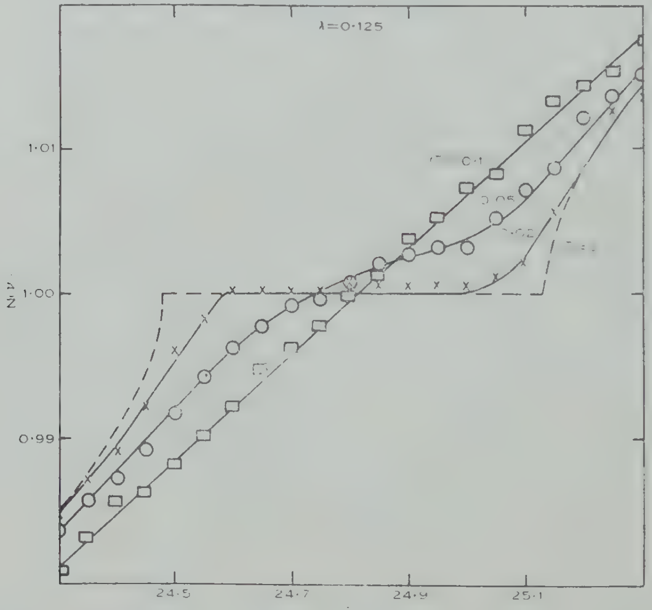


Fig. 1.—The number of states per length a with energy E less than that given by $\nu \propto E^{\frac{1}{2}}$, for increasing disorder described by the parameter σ. The value $\sigma=0$ refers to the ordered model of Kronig and Penney. The region near the first forbidden zone is shown for a "weak binding" situation.

The results differ from those of the periodic (Kronig-Penney) model only in the vicinity of forbidden zones, as might be expected, and are therefore shown only for these ranges of ν. For $\lambda=0\cdot125$ (a weak-binding case, Fig. 1), the first forbidden zone is shown. With $\lambda=2$, a first forbidden zone occurs around $\nu=0$; this is a "tight-binding" situation, the missing positive energy states going to fill a negative energy band. Examination of this region has been deferred till the negative energy states are treated ; hence the results presented, Figure 2, for $\lambda=2$ are near the second forbidden zone.

(*b*) *The Effect of Short-range Disorder*

Both short- and long-range disorder occur in the chains considered above. Some light is thrown on their relative importance in narrowing the forbidden band by the results shown in Figure 3. The potentials $V(x)$ used for this were obtained from the particular chains used in Figures 1 and 2 by replacing each

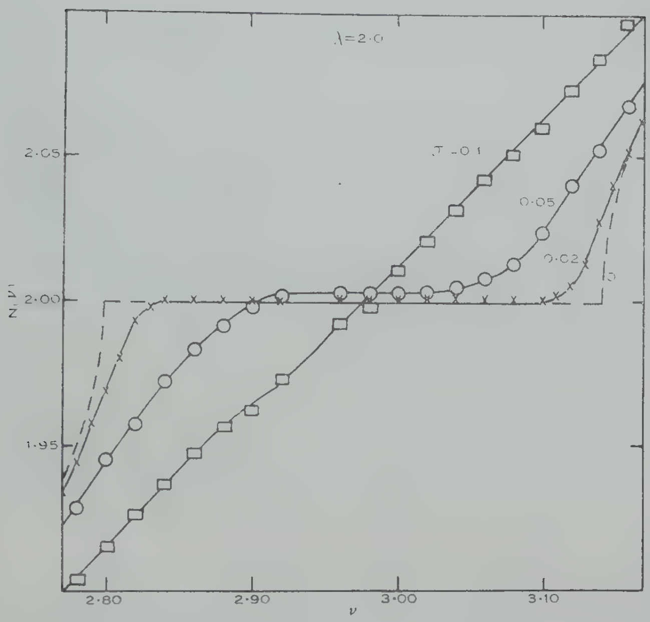


Fig. 2.—As in Figure 1 but for a much higher value of the parameter λ corresponding to "tight binding". The second forbidden zone is shown.

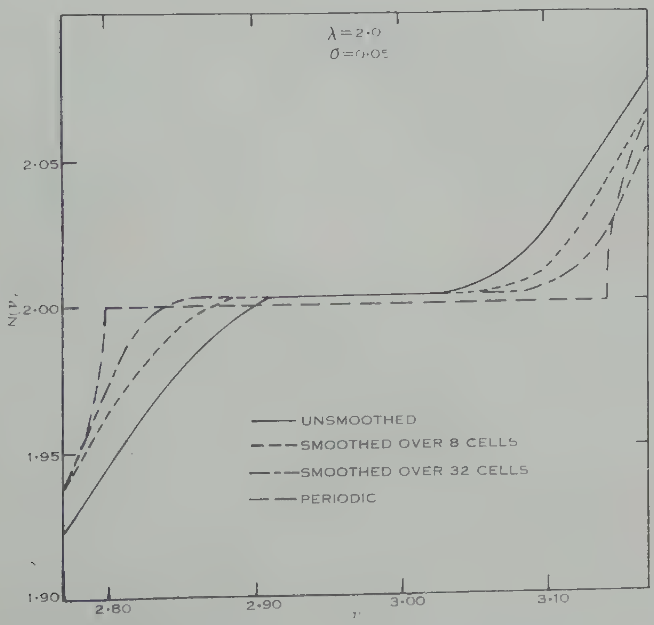


Fig. 3.—Showing the effect on the width of the forbidden zone of "smoothing out" the short-range disorder.

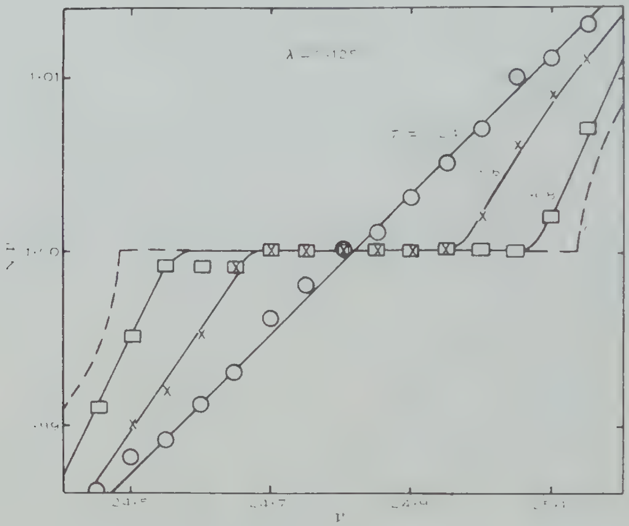


Fig. 4.—Long-range order preserved, weak binding ; increase of the parameter τ corresponds to increasing short-range disorder. The first forbidden zone.

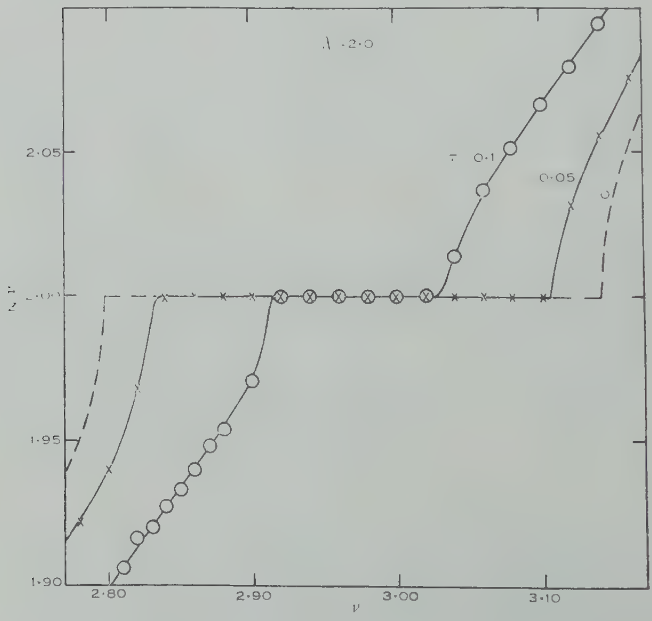


Fig. 5.—As in Figure 4 but for tight binding. The second forbidden zone.

u_n by the average u over s cells of the original chain adjoining the nth. Short-range order can thus be progressively restored by increasing s leaving the violation of long-range order substantially unaffected. Curves thus " smoothed " with $s=0$, 8, and 32 are shown in Figure 3.

A different set of probability distributions was used to obtain the results for short-range disorder alone, Figures 4 and 5. Here long-range order was preserved by choosing x_n to be either $na+\tau$ or $na-\tau$, each with probability $\frac{1}{2}$. Initial test runs showed no fluctuation whatsoever between the results for chains of 1000 and 2000 cells respectively, and hence chains of 1000 cells were chosen to give the displayed results.

The possibility was considered that this " single step size " distribution might give qualitatively different results from those for a smooth continuous distribution. As a check, a similar calculation was performed for the case of long-range disorder, letting $u=1\pm\tau$ with $\tau=0\cdot06\surd 5/8$. The mean deviation of u then equalled that of the continuous distribution described above, for $\sigma=0\cdot02$. Comparison of the two sets of results showed unimportant differences.

III. Discussion

In interpreting the figures, it should be noted that a horizontal portion of a curve for $N(\nu)$ signifies an absence of states in that range of ν, i.e. an energy gap. The $N(\nu)$ curve for free electrons would be a straight line of slope λ/π; the curves for $\sigma=0\cdot1$ in Figures 1 and 2 have approximately this slope.

The curves for $\sigma=0\cdot02$ in Figures 1 and 2 show no increase, to within the node-count error of ±1, in the number of states as the energy increases through the range where a gap may be expected. We take this as supporting the conclusion that in an infinite chain there would be indeed a forbidden zone for values of σ not too large. For $\sigma=0\cdot02$ the gap is narrowed but still apparent, while for $\sigma=0\cdot05$ it is reduced to less than half of its periodic value (for $\lambda=2$) or practically absent ($\lambda=0\cdot125$). Thus, for potential strengths varying by a factor of 16, most of the gap has been occupied for some σ_c satisfying $0\cdot02<\sigma_c<0\cdot05$.

This corresponds to the " atoms " being half a lattice length away, on the average, from ordered positions after 100 to 600 lattice lengths.

In comparing the results with the conclusions of Gubanov (1954) the following point is relevant. Gubanov applies a transformation $x\rightarrow\xi$ in which $\xi(x_n)=na$, and treats the resulting potential $V(\xi)$ as periodic. He points out that this involves the neglect of variations in the shape of the potential in different cells. Study of Figures 3, 4, and 5 may raise doubt as to the validity of this neglect of short-range disorder in the case of a general potential. However, for the special potential considered here, a delta-function, and for this potential only, these variations are absent, and his transformation gives an accurately periodic lattice. This suggests that the present model and calculations constitute a fair test of his theory.

In so far as the results show a definite gap, narrowed as σ increases, they agree with Gubanov's conclusion. However, in two important respects there is disagreement.

In Gubanov's theory, all the energy levels of the periodic model are shifted upward, and the gap-narrowing is a consequence simply of a difference in the magnitude of the upward shifts of the levels bounding the gap. However, it may be seen, Figures 1 and 2, that in the model chosen here there is indeed an upward shift of the lower gap edge, but a downward shift of the upper gap edge contrary to Gubanov's conclusion.

Also, while Gubanov gives no criterion for preservation of the gap, he makes an order of magnitude estimate of a representative narrowing as about 0·1 eV for $\sigma=0\cdot1$. In our model, the gap is totally destroyed for this value of σ.

Figures 3, 4, and 5 indicate that, in one dimension, narrowing of the energy gap is mainly brought about by lack of short-range order (apparent within say 30 "interatomic" distances), whether or not there is order over longer ranges. The only apparent effect of the latter is on the sharpness of the corner in curves as they approach the gap.

A smooth gradual distortion of a regular chain, so that it loses long-range order over say 100 lattice lengths, evidently produces remarkably little alteration in the energy bands.

A given narrowing of the gap, say to half, may be produced either by disorder (long and short) of the type specified in Section II (*a*), or by short-range disorder alone such as specified in II (*b*).

In conclusion, some points may be noted concerning the relevance of these calculations to the energy band problem for a real, three-dimensional liquid. One realizes that a one-dimensional model cannot be expected to reproduce those aspects of the behaviour of a real liquid dependent on interatomic forces ; however, the object of the present discussion is not to discuss such properties but only the effect of a type of atomic disorder on the electrons concerned in conduction.

X-ray diffraction experiments show (e.g. Gingrich 1943) that 0·15 is a representative figure for the r.m.s. fluctuation in the nearest-neighbour separation of some liquid alkali metals near the melting point. If we identified this parameter with our σ and supposed the results applied in three dimensions, we would conclude quite wrongly that the gap must be eliminated on melting.

However, as we have seen, in one dimension the band structure is disturbed primarily by short-range disorder. In three dimensions, however, the scattering centres are distributed over a whole wave front, and we would intuitively expect the disorganizing effect of their short-range disorder on the electron wave function to be much less than in the one-dimensional case.

Further, the uncertainty at present as to what is a suitable model for a real liquid (Bernal 1959, cf. Gubanov 1955) makes any quantitative conclusions hazardous.

IV. REFERENCES

BERNAL, J. D. (1959).—*Nature* **183**: 141.
EISENSCHITZ, R., and DEAN, P. (1957).—*Proc. Phys. Soc. Lond.* **70**: 713.
GINGRICH, N. S. (1943).—*Rev. Mod. Phys.* **15**: 90.
GUBANOV, A. I. (1954).—*Zh. Eksp. Teor. Fiz.* **26**: 139.
GUBANOV, A. I. (1955).—*Zh. Eksp. Teor. Fiz.* **28**: 401.
KITTEL, C. (1956).—" Introduction to Solid State Physics." (Wiley: New York.)
KRONIG, R., and PENNEY, W. G. (1931).—*Proc. Roy. Soc.* A **130**: 499.
LANDAUER, R., and HELLAND, J. C. (1954).—*J. Chem. Phys.* **22**: 1655.
LAX, M., and PHILLIPS, J. C. (1958).—*Phys. Rev.* **110**: 41.
MOTT, N. F., and JONES, H. (1936).—" The Theory of the Properties of Metals and Alloys." (Oxford Univ. Press.)
ROJANSKY, V. (1938).—" Introductory Quantum Mechanics.' (Prentice-Hall: New York.)

ALTERNATIVE SUPERCONDUCTING GROUND STATES

By J. C. Fisher*

[*Manuscript received October 8, 1959*]

Summary

A comparison is made between (*a*) the trial ground state composed of a linear combination of normal state configurations in which individual electron states of opposite momentum and equal spin projection ($k\uparrow$, $-k\uparrow$) and ($k\downarrow$, $-k\downarrow$), abbreviated ($k\sigma$, $-k\sigma$), are both either occupied or unoccupied, and (*b*) the trial ground state with ($k\uparrow$, $-k\downarrow$) pairing proposed by Bardeen, Cooper, and Schrieffer. The two trial states lead to identical thermodynamics as long as an average matrix element $-V=\langle V'_{kk'}\rangle_{\mathrm{av}}$ is treated as a disposable parameter. They lead to similar coherence effects for ultrasonic attenuation and nuclear spin relaxation. They differ in that spin paramagnetism is predicted for ($k\sigma$, $-k\sigma$) pairing. Detailed knowledge of the electron-phonon interaction in superconductors seems at present inadequate to rule out either possibility for all superconductors.

I. Introduction

In constructing the ground state wave function for a superconductor, Bardeen, Cooper, and Schrieffer (1957) took a linear combination of normal state configurations in which individual electron states of equal but opposite momentum and spin projection ($k\uparrow$, $-k\downarrow$) are both either occupied or unoccupied,

$$\Phi_0=\Pi_k(1+g_k^2)^{-\frac{1}{2}}(1+g_kC^*_{k\uparrow}C^*_{-k\downarrow})\mid 0\rangle. \quad \ldots\ldots\ldots\ldots \quad (\mathbf{1a})$$

The major properties of superconductors have been derived from this choice of ground state, with the exception of the spin paramagnetism observed in superconducting small particles by Reif (1957) and in thin films by Androes and Knight (1959).

Valatin (1958) noted that $(C_{k\uparrow}-g_kC^*_{-k\downarrow})\Phi_0=0$ for all k, and defined the new variables

$$\left.\begin{aligned}\xi_\varkappa&=(1+g_\varkappa^2)^{-\frac{1}{2}}(C_\varkappa-g_\varkappa C^*_{-\varkappa}),\\ \xi^*_\varkappa&=(1+g_\varkappa^2)^{-\frac{1}{2}}(C^*_\varkappa-g_\varkappa C_{-\varkappa}),\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots \quad (2)$$

with

$$\left.\begin{aligned}\varkappa&=k,\ \sigma,\\ -\varkappa&=-k,\ -\sigma,\\ g_\varkappa&=(\mathrm{sig}\ \sigma)g_k.\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{3a})$$

The ξ's satisfy the commutation relationships for Fermi excitations, and the transformation from the C's to the ξ's is canonical. The ground state can be written

$$\Phi_0=\Pi_{\varkappa\neq-\varkappa}(1+g_\varkappa^2)^{-\frac{1}{2}}(1+g_\varkappa C^*_\varkappa C^*_{-\varkappa})\mid 0\rangle, \quad \ldots\ldots\ldots\ldots \quad (4)$$

---

* General Electric Research Laboratory, Schenectady, New York, U.S.A.

where the product runs over half the values of $\varkappa$, including $\varkappa$ or $-\varkappa$ but not both. It is defined by the relationships

$$\xi_\varkappa \Phi_0 = 0 \qquad \text{(5)}$$

for all $\varkappa$, and it serves as vacuum state for the complete set of quasi-particle excitations

$$\xi^*_{\varkappa_1} \xi^*_{\varkappa_2} \ldots \xi^*_{\varkappa_j} \Phi_0. \qquad \text{(6)}$$

Equations (2), (4), (5), (6), giving the canonical transformation, the ground state Φ_0, the identification of Φ_0 as the vacuum state for quasi-particles in the ξ representation, and the excited states in the ξ representation, all are written in terms of the index $\varkappa$ and the parameter $g_\varkappa$ defined in equation (**3a**). If the definitions of $\varkappa$ and $g_\varkappa$ are altered to read

$$\left.\begin{aligned} \varkappa &= k,\ \sigma, \\ -\varkappa &= -k,\ \sigma, \\ g_\varkappa &= (\operatorname{sig} k_z) g_k, \end{aligned}\right\} \qquad \text{(3b)}$$

equations (2), (4), (5), (6) still hold, although when written in terms of k and σ the ground state is seen to consist of a linear combination of normal state configurations involving electron pairs with *equal* spin projections $(k\uparrow,\ -k\uparrow)$ and $(k\downarrow,\ -k\downarrow)$, abbreviated $(k\sigma,\ -k\sigma)$:

$$\Phi'_0 = \Pi_{\sigma,\, k(k_z>0)} (1+g_k^2)^{-\frac{1}{2}} (1 + g_k C^*_{k\sigma} C^*_{-k\sigma}) \mid 0\rangle. \qquad \text{(1b)}$$

The trial ground state Φ'_0 with $(k\sigma,\ -k\sigma)$ pairing is of interest because it leads to spin paramagnetism in the superconducting phase, whereas Φ_0 with $(k\uparrow,\ -k\downarrow)$ pairing does not. The distinction comes about because with $(k\sigma,\ -k\sigma)$ pairing, an integral number of pairs can be flipped from spin down to spin up, giving a net spin polarization, without breaking any pairs. A continuation of Valatin's analysis clarifies this point and at the same time provides a comparison of the two types of pairing from other standpoints.

II. Comparison of ↑↓ and σσ Pairing

First consider the matrix elements of a single-particle scattering operator $B = \Sigma_{kk'\sigma\sigma'} B_{kk'} C^*_{k\sigma} C_{k'\sigma'}$ between pairs of excited states of a superconductor. The scattering operator can be written in a form that is valid for either definition of the index $\varkappa$,

$$\left.\begin{aligned} B &= \Sigma_{\varkappa\varkappa'} B_{kk'} C^*_\varkappa C_{\varkappa'} \\ &= \Sigma (B_{kk'} C^*_\varkappa C_{\varkappa'} + B_{-k'-k} C^*_{-\varkappa'} C_{-\varkappa}) \\ &= \Sigma B_{kk'} (C^*_\varkappa C_{\varkappa'} \pm C^*_{-\varkappa'} C_{-\varkappa}). \end{aligned}\right\} \qquad \text{(7)}$$

Terms of the form $C^*_\varkappa C_{\varkappa'}$ and $C^*_{-\varkappa'} C_{-\varkappa}$ are grouped together in the sum because they prove to be coherent.

The matrix elements appropriate to ultrasonic attenuation are of the form

$$(C^*_{k_1\uparrow} C_{k_2\uparrow} + C^*_{-k_2\downarrow} C_{-k_1\downarrow}) \qquad \text{(8a)}$$

for $(k\uparrow,\ -k\downarrow)$ pairing, and

$$(C^*_{k_1\sigma} C_{k_2\sigma} + C^*_{-k_2\sigma} C_{-k_1\sigma}) \qquad \text{(8b)}$$

for $(k\sigma, -k\sigma)$ pairing. The nature of the electron-scattering process is such that there is no spin-flip ($C^*_\uparrow C_\downarrow$ terms do not appear), and there is practically no k_z-flip (the scattering is very small angle at the ultrasonic frequencies so far employed). The absence of spin-flip terms for ↑↓ pairing in (**8a**) leads to certain coherence effects, and the absence of k_z-flip for $\sigma\sigma$ pairing leads to the same ones.

The matrix elements appropriate to nuclear spin relaxation, on the other hand, are of the form

$$\left.\begin{array}{l} (C^*_{k_1\uparrow}C_{k_2\uparrow}-C^*_{-k_2\downarrow}C_{-k_1\downarrow}) \\ (C^*_{k_1\uparrow}C_{-k_2\downarrow}+C^*_{k_2\uparrow}C_{-k_1\downarrow}) \\ (C^*_{k_1\downarrow}C_{-k_2\uparrow}+C^*_{k_2\downarrow}C_{-k_1\uparrow}) \end{array}\right\} \quad \ldots\ldots\ldots\ldots \quad (\mathbf{9a})$$

for ↑↓ pairing, and

$$\left.\begin{array}{l} (C^*_{k_1\sigma}C_{k_2\sigma}+C^*_{-k_2\sigma}C_{-k_1\sigma}) \\ (C^*_{k_1\sigma}C_{-k_2-\sigma}+C^*_{k_2-\sigma}C_{-k_1\sigma}) \end{array}\right\} \quad \ldots\ldots\ldots\ldots\ldots \quad (\mathbf{9b})$$

for $\sigma\sigma$ pairing. The coherence effects for ↑↓ pairing are such that each squared matrix element contains the factor

$$[(hh')^{\frac{1}{2}}+((1-h)(1-h'))^{\frac{1}{2}}]^2,$$

where $h=g^2_{k_1}(1+g^2_{k_1})^{-1}$, $h'=g^2_{k_2}(1+g^2_{k_2})^{-1}$, as shown by Hebel and Slichter (1959). For $\sigma\sigma$ pairing these coherence effects are spoiled, the sign of one part of the matrix element relative to the other depending upon the presence or absence of a k_z-flip. Since k_1, k_2 pairs have k_{1z} and k_{2z} each positive half the time and negative half the time, each squared matrix element contains a factor

$$[(hh')^{\frac{1}{2}}\pm((1-h)(1-h'))^{\frac{1}{2}}]^2$$

where the ± sign is a random variable. Since half the squared matrix elements have the positive sign (agreeing with Hebel and Slichter's result) and half have the negative sign (giving the same result as for ultrasonic attenuation), the curve of relaxation rate versus temperature corresponding to the mixture of ± signs should be the average of (*a*) one of the curves calculated by Hebel and Slichter for different assumed breadths of the electron spin energy levels and (*b*) the ultrasonic attenuation curve. The best of these curves agrees well with experiment, rising sharply just below T_c, then dropping again far below T_c.

Next consider the ground and excited state energies. Following Valatin, the Hamiltonian

$$H=E_0+\Sigma_{k\sigma}\varepsilon_k C^*_{k\sigma}C_{k\sigma}+(1/2)\Sigma_{kk'q\sigma\sigma'}V_{kk'}C^*_{k\sigma}C^*_{q-k,\sigma'}C^*_{q-k',\sigma'}C_{k'\sigma} \quad \ldots \quad (\mathbf{8})$$

is transformed to the ξ representation by the inverse of equations (**2**). The portion of the transformed Hamiltonian that is diagonal in the states $\xi^*_{\varkappa_1}\xi^*_{\varkappa_2}\ldots\xi^*_{\varkappa_j}\Phi_0$ is of the form

$$H_0=W_0+\Sigma_\varkappa W_k N_\varkappa+\Sigma_{\varkappa\neq\varkappa'}W_{\varkappa\varkappa'}N_\varkappa N_{\varkappa'}, \quad \ldots\ldots\ldots \quad (\mathbf{9})$$

where $N_\varkappa=\xi^*_\varkappa\xi_\varkappa$ is the quasi-particle number operator, and where W_0 and W_k are respectively the ground state energy and the excitation energy for the states $\xi^*_{k\uparrow}\Phi_0$ and $\xi^*_{k\downarrow}\Phi_0$. Both W_0 and W_k are invariant in functional form with respect to the two definitions of $\varkappa$ and $g_\varkappa$, when the matrix element $V'_{kk'}$ (shortly to be defined) is substituted for $V_{kk'}$ in [the parameter μ_k appearing in] Valatin's expressions for W_0 and W_k. The only difference attributable to the choice of $(k\uparrow, -k\downarrow)$ or $(k\sigma, -k\sigma)$ pairing is contained in this matrix element $V'_{kk'}$, which has the values

$$V'_{kk'} = V_{kk'} \quad \dots\dots\dots\dots\dots\dots\dots\dots \quad \mathbf{(10a)}$$

for $(k\uparrow, -k\downarrow)$ pairing, and

$$V'_{kk'} = V_{kk'} \operatorname{sig} k_z k'_z \quad \dots\dots\dots\dots\dots\dots \quad \mathbf{(10b)}$$

for $(k\sigma, -k\sigma)$ pairing. Even this difference disappears when $V'_{kk'}$ is replaced by an average matrix element $-V$ over the energy range $-\hbar\omega < \varepsilon_k < \hbar\omega$, as in Bardeen, Cooper, and Schrieffer's procedure for evaluating and minimizing W_0. The two trial functions then yield identical results throughout.

When $(k\sigma, -k\sigma)$ pairing is employed, the energy of the superconducting state can be written

$$W_0 = -\Sigma_\sigma N_{0\sigma}(\hbar\omega_\sigma)^2 \left[\exp(2/N_{0\sigma}V_\sigma) - 1\right]^{-1}, \quad \dots\dots \quad \mathbf{(11)}$$

where $N_{0\uparrow}$ need not equal $N_{0\downarrow}$. Here W_0 is measured with respect to the normal state with the same values of $N_{0\sigma}$. The cut-off frequency ω_σ and the average matrix element $-V_\sigma$ both may vary slowly with $N_{0\sigma}$. Since W_0 varies quadratically with spin polarization $(N_{0\uparrow} - N_{0\downarrow})$, spin paramagnetism is to be expected wherever a magnetic field penetrates the superconductor. It is difficult to estimate the magnitude of the susceptibility that would be expected in the superconducting phase, because of the possible variation of V_σ with $N_{0\sigma}$. If V_σ were constant, the susceptibility of the superconducting phase would exceed that of the normal phase.

As long as $-V = \langle V'_{kk'}\rangle_{\text{av}}$ is taken as a disposable parameter, the question as to which pairing leads to the lower ground state energy cannot be settled, since identical results then are obtained throughout (except for spin paramagnetism in the superconducting phase which has not as yet been accounted for by $(k\uparrow, -k\downarrow)$ pairing, although Heine and Pippard (1958) have suggested how it might be done). A direct estimate of $\langle V'_{kk'}\rangle_{\text{av}}$ is therefore in order for the two types of pairing.

Physically, the relative magnitudes of $\langle V_{kk'}\rangle_{\text{av}}$ and of

$$\langle V'_{kk'}\rangle_{\text{av}} = \langle V_{kk'} \operatorname{sig} k_z k'_z\rangle_{\text{av}}$$

will depend upon the relative ranges of the phonon-induced attraction that is responsible for superconductivity and the screened Coulomb repulsion that opposes it. Both are reduced in magnitude for $\sigma\sigma$ pairing relative to $\uparrow\downarrow$ pairing, but if the screened Coulomb repulsion is of sufficiently short range, it will be reduced much more, causing $\sigma\sigma$ pairing to give the stronger net attraction. Numerical calculation is difficult, although Pines (1958) has made perhaps the best estimate of $V_{kk'}$ possible at the present time. His result suggests that $\uparrow\downarrow$ pairing can sometimes lead to superconductivity, but $\sigma\sigma$ pairing seems less likely to do so.

III. CONCLUSIONS

Since the coherence effects associated with $(k\uparrow, -k\downarrow)$ and $(k\sigma, -k\sigma)$ pairing are in equally good agreement with ultrasonic attenuation and nuclear spin relaxation measurements, and since our knowledge of the details of the electron-phonon interaction in superconductors is at best approximate, perhaps the best conclusion that can be reached is that neither the $\uparrow\downarrow$ nor the $\sigma\sigma$ configuration can be ruled out completely at present.

IV. ACKNOWLEDGMENTS

The author wishes to thank Dr. M. H. Cohen, Dr. W. A. Harrison, Dr. E. W. Hart, Dr. V. Heine, and Dr. T. O. Woodruff for a number of stimulating discussions.

V. REFERENCES

ANDROES, G. M., and KNIGHT, W. D. (1959).—*Phys. Rev. Letters* **2**: 386.
BARDEEN, J., COOPER, L. N., and SCHRIEFFER, J. R. (1957).—*Phys. Rev.* **108**: 1175.
HEBEL, L. C., and SCHLICHTER, C. P. (1959).—*Phys. Rev.* **113**: 1504.
HEINE, V., and PIPPARD, A. B. (1958).—*Phil. Mag.* **3**: 1046.
PINES, D. (1958).—*Phys. Rev.* **109**: 280.
REIF, F. (1957).—*Phys. Rev.* **106**: 208.
VALATIN, J. G. (1958).—*Nuovo Cim.* **7**: 843.

SUSCEPTIBILITY OF CHROMIUM–BASED ALLOYS OF TRANSITION ELEMENTS

By W. M. Lomer*

[*Manuscript received September* 25, 1959]

Summary

The magnetic properties of chromium-based alloys with the transition metals are very varied. Below room temperature it is believed that chromium is antiferromagnetic and in this paper we investigate briefly the nature of the influence of this antiferromagnetic matrix on the properties of dilute solutions of atoms carrying well-defined local moments of their own. It is shown that various anomalies in the susceptibility-temperature relations can be rationalized in terms of a model involving small spin orientation forces arising from the antiferromagnetism. Alloying with vanadium appears to reduce the spin moment on the chromium atoms, but to introduce no localized spins itself. The temperature-independent contribution of cobalt in solution is interpreted in terms of local spins locked by the antiferromagnetism of the matrix. The results so far obtained suggest that careful specific heat measurements in the range 0–100 °K may disclose very significant anomalies in many of these alloys.

I. Introduction

The electronic structure of pure chromium has been much debated. The occurrence of a small sharp maximum in the susceptibility curve has been detected at or near 300 °K by Wagenknecht (1956) and by Lingelbach (1958). The occurrence at or near the same temperature of anomalies in several other properties has been reported (see e.g. Sully 1954). The elastic constants, electrical resistivity, and lattice constant show sudden changes in their temperature coefficients but no discontinuity in their values. No specific heat anomaly has been observed, and indeed there is still room for doubt as to whether all of the anomalies occur at the same temperature in the same specimen. It was early suggested that these effects were the manifestation of antiferromagnetism in the metal, though at first sight it appears as if the absence of a specific heat anomaly should rule this out. Great interest has therefore attached to neutron diffraction from pure chromium.

The first experiment reported was that of Shull and Wilkinson (1953), who found distinct evidence of a magnetic superlattice. The strength of the diffraction was consistent with a spin density of about $0 \cdot 2$–$0 \cdot 4$ Bohr magnetons per atom. They reported a Neel temperature of 450 °K, well above the usual anomaly temperatures. More recently, unpublished work by Weiss, Corliss, and Hastings at Brookhaven National Laboratory on single crystals, and by Bacon at Harwell on de-oxidized chromium produced by Aeronautical Research Laboratories, Melbourne, Australia, have shown that the situation is far from simple. The results for the Neel temperature and the domain structure of the antiferro-

---

*Atomic Energy Research Establishment, Harwell, Berks., England.

magnetism appear to vary with the purity and with the single or polycrystallinity of the metal. It is clearly premature to attempt to deduce much from this work at this time except that chromium is often antiferromagnetic at low temperatures and that alloying can change the critical temperature. The work also serves as a warning that large discrepancies may arise between samples of different origin or heat treatment.

The spread of temperatures for various anomalies therefore needs most careful attention, for elasticity, resistivity, and susceptibility experiments have not normally been carried out on one and the same sample.

The absence of the specific heat anomaly is probably connected with the absence of a true Curie-Weiss behaviour in pure chromium at high temperatures. It is likely that the transition is one from an ordered spin arrangement of *d*-electrons to a disordered array in which the system is effectively a degenerate electron gas, in which to align many of the spins parallel would need considerable energy so that only a Pauli-type paramagnetism would occur. The theory by Lidiard (1954) is of this nature. Since the entropy of the degenerate electron gas is low there will not be any considerable specific heat anomaly. The suggestion is therefore that there are two states of nearly zero total spin ; one with spin density zero everywhere, and one with a periodic superlattice of opposite spins. These states must have nearly equal internal energy in order that the very small entropy difference can cause the transition. The smallness of these terms is probably the key to the purity and structure sensitivity of the system.

II. LOCALIZED IMPURITY SPINS IN AN ANTIFERROMAGNETIC LATTICE

To get a rough idea of the nature of the behaviour of isolated impurity spins in an antiferromagnetic lattice, we consider the following ideal situation. About any impurity lattice site we suppose that the spins of the neighbour atoms are of a particular sign, and that there are exchange forces of unspecified origin which give rise to an exchange energy splitting 2Δ between the two spin orientations of the impurity atom (assumed to be $S=\frac{1}{2}$). We further assume that there are no interactions between the impurity spins and that the populations of spin-up sites and spin-down sites are equal at all times to $\frac{1}{2}N$ on each type.

In the presence of a magnetic field parallel to the spin lattice therefore, the energy levels for spin-up sites $\pm(\Delta+\mu H)$ and for spin-down sites are $\pm(\Delta-\mu H)$ where μ is the Bohr magneton. The partition functions for these two types of sites are:

$$Z \text{ up},\ 2\cosh(\Delta+\mu H),$$
$$Z \text{ down},\ 2\cosh(\Delta-\mu H).$$

The susceptibility χ is given by $N\mu\partial^2(kT\ln Z)/\partial H^2$ and the sum of the susceptibilities for the two types of site is:

$$\chi=\frac{N}{2}\,\frac{\mu^2}{kT}\left[\frac{1}{\cosh^2\{(\Delta+\mu H)/kT\}}+\frac{1}{\cosh^2\{(\Delta-\mu H)/kT\}}\right],$$

and

$$\chi_{H\to 0}=\frac{N\mu^2}{kT}\,\mathrm{sech}^2\,\frac{\Delta}{kT}.$$

For $kT \gg \Delta$ the susceptibility is simply the Curie expression $N\mu^2/kT$ (note that the mean square magnetic moment $m^2 = 3\mu^2$ and χ is therefore equal to $m^2N/3kT$), and when $\Delta \gg kT$ it becomes $(4N\mu^2/kT) \exp(-2\Delta/kT)$. Thus once kT decreases below Δ the parallel susceptibility is rapidly eliminated (see Fig. 1).

In the presence of a magnetic field perpendicular to the spin lattice, the energy levels may be found as the solution of a simple two by two secular equation. The matrix element connecting the spin-up to spin-down sites is of magnitude μH, and the eigenvalues of the energy are given as the roots of

$$\begin{vmatrix} -\Delta - E & \mu H \\ \mu H & \Delta - E \end{vmatrix} = 0,$$

so that $E = \pm(\Delta^2 + \mu^2 H^2)^{\frac{1}{2}}$.

There is no distinction between the two types of site this time and

$$Z = 2 \cosh (\Delta^2 + \mu^2 H^2/kT)^{\frac{1}{2}},$$

$$M = \frac{kT}{Z}\,\frac{\mathrm{d}Z}{\mathrm{d}H} = \frac{N\mu^2 H}{(\Delta^2 + \mu^2 H^2)^{\frac{1}{2}}} \tanh (\Delta^2 + \mu^2 H^2/kT)^{\frac{1}{2}},$$

$$\chi_{H\to 0} = (N\mu^2/\Delta) \tanh \Delta/kT.$$

Thus for small Δ/kT, we regain the Curie law, while for large Δ/kT, the susceptibility tends to the constant value $N\mu^2/\Delta$.

For a polycrystalline medium the mean susceptibility will be $\frac{2}{3}\chi_\perp + \frac{1}{3}\chi_{||}$ and is therefore

$$\frac{N\mu^2}{3}\left\{\frac{1}{kT} \operatorname{sech}^2 \frac{\Delta}{kT}\right\},$$

and this is plotted in Figure 1.

Now the energy gap Δ may be taken, in a crude phenomenological way, to the degree of saturation in the antiferromagnetic array and is therefore expected to vary roughly as a Brillouin curve with temperature. This implies that at $0{\cdot}9T_c$ Δ is very nearly as great as Δ at absolute zero. There is, however, no obvious necessary connexion between T_c and the magnitude of Δ, at least for dilute alloys.

It is possible for Δ to be considerably smaller than T_c, and under these circumstances we may find that, as the material is cooled, the contribution to the susceptibility attributable to the dilute magnetic atoms remains true Curie (or Curie-Weiss when interactions are included) until the Neel point is reached; (at 3Δ in Fig. 2). At this point a rapid fall in the susceptibility ensues as Δ builds up to its full value, and the value for small Δ/kT is $\chi_{\text{curie}}\left\{1 - \frac{7}{9}\cdot(\Delta/kT)^2\right\}$. Thereafter both $\chi_{||}$ and $\chi_\perp$ continue to increase roughly as $1/T$ until $kT \sim \Delta$, when another change of shape results. A maximum susceptibility followed by a constant region as $T \to 0$ is the result of the suppression of $\chi_{||}$ and the levelling of $\chi_\perp$ to its value $N\mu^2/\Delta$.

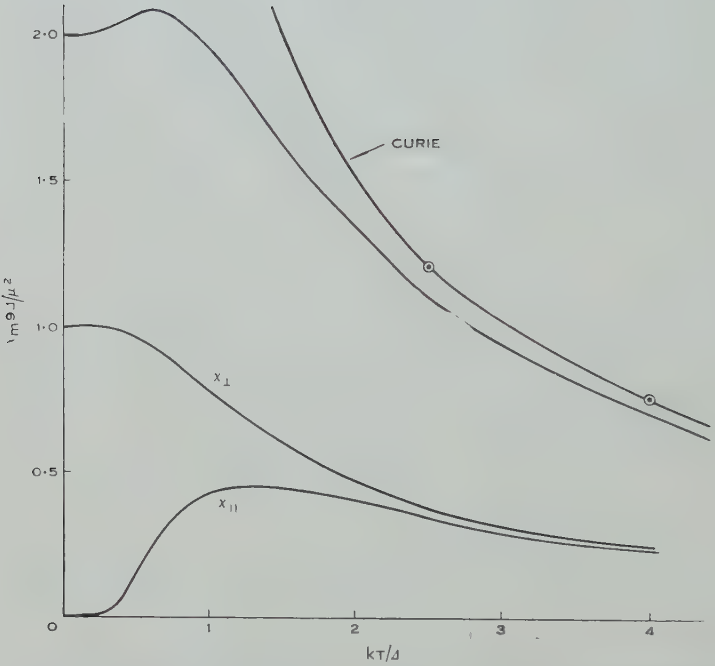


Fig. 1.—Calculated curves of $\chi_{||}$, $\chi_\perp$, and $\chi_m=\frac{1}{3}\chi_{||}+\frac{2}{3}\chi_\perp$ for two spin levels of energy separation 2Δ compared with the simple Curie $1/T$ relation.

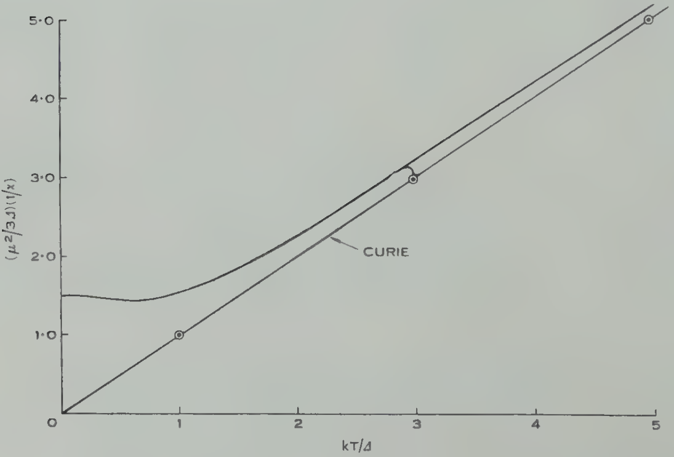


Fig. 2.—Calculated curve of $1/\chi_m$ compared with simple Curie law.

If Δ is larger than kT_c then a temperature-independent contribution from χ is all that may be expected at temperatures where the antiferromagnetic structure is fully established. Since there is no relation between Δ and T_c, the region in which we side step from the lower to the upper curve in Figure 2 is unknown.

III. EXPERIMENTAL RESULTS

Three recent sets of experiments seem to support the general picture presented above. The most direct are those of Newmann and Stevens (1959), which were carried out on binary alloys of iron in chromium. For a series of compositions, including 1 and 2 atomic per cent. as well as higher concentrations, susceptibility, resistivity, and coefficient of expansion were measured over a range of temperatures from 80 to 600 °K. Anomalies were found in all of these properties at about 300 °K, with a tendency for this temperature to fall as the concentration of iron rose (see Fig. 3). In particular the susceptibility due to the iron additions, though it increased very roughly as $1/T$ over the whole range, dropped by about 4 or 5 per cent. as the temperature fell through the anomalous region (see Fig. 4). The small Δ/kT expansion of the previous section shows that this corresponds to Δ/kT_c of about $0\cdot23$, so that a levelling off in the susceptibility would be expected to show up at about $0\cdot23\times300$ °K$\sim$70 °K. There is no sign of this levelling off at 80 °K, but lower temperature measurements would test the present hypothesis. The suggestion by Stevens and Newmann that the drop in susceptibility and other anomalies are due to a change in the antiferromagnetic structure rather than its first appearance (to reconcile their results with those of Shull and Wilkinson) should have the same consequences at low temperatures as our hypothesis. The work of de Vries (1959) cited later (see Fig. 5) seems to make it very likely that the first appearance of antiferromagnetic structure is strongly composition dependent and makes the need to reconcile the Shull and Wilkinson results and those of Newmann and Stevens less acute, and the assumption of alternative spin structures less attractive. Only careful neutron diffraction experiments could distinguish between the two alternatives.

Lingelbach carried out a series of measurements on susceptibility of several chromium-based binary and ternary alloys at temperatures between 80 and 300 °K. His curves fall sharply into two classes, those with temperature-independent additional magnetization, and those with approximate Curie-Weiss behaviour. Several have been collected together and redrawn in Figure 6. Among the former are the cobalt and vanadium binary alloys, and among the latter the iron binary alloys. However, vanadium has a comparatively small effect, for 10% V has only about two-thirds of the effect of $1\cdot5$% Co. It seems plausible to suggest that the cobalt contribution is temperature independent because the large atomic cobalt moment is pinned down, while vanadium carries no atomic moment of its own, and modifies only the Pauli free electron spin paramagnetism of the chromium. The temperature-dependent susceptibilities we attribute to incompletely locked spins, and the behaviour of the iron binary alloys is consistent with the observation of Stevens and Newmann, though the temperature range studied by Lingelbach does not include the anomaly.

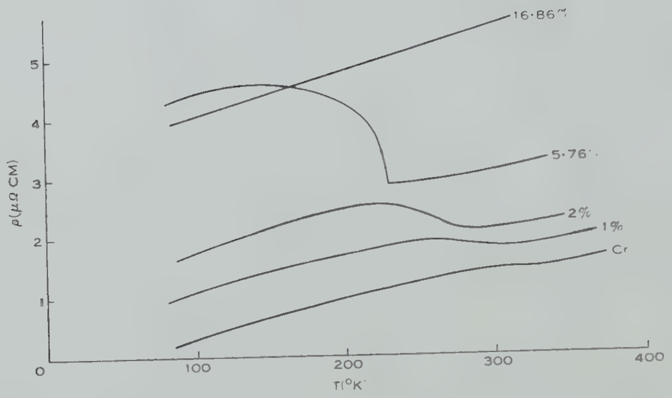


Fig. 3.—Resistivity of iron alloys of various compositions (from Newmann and Stevens).

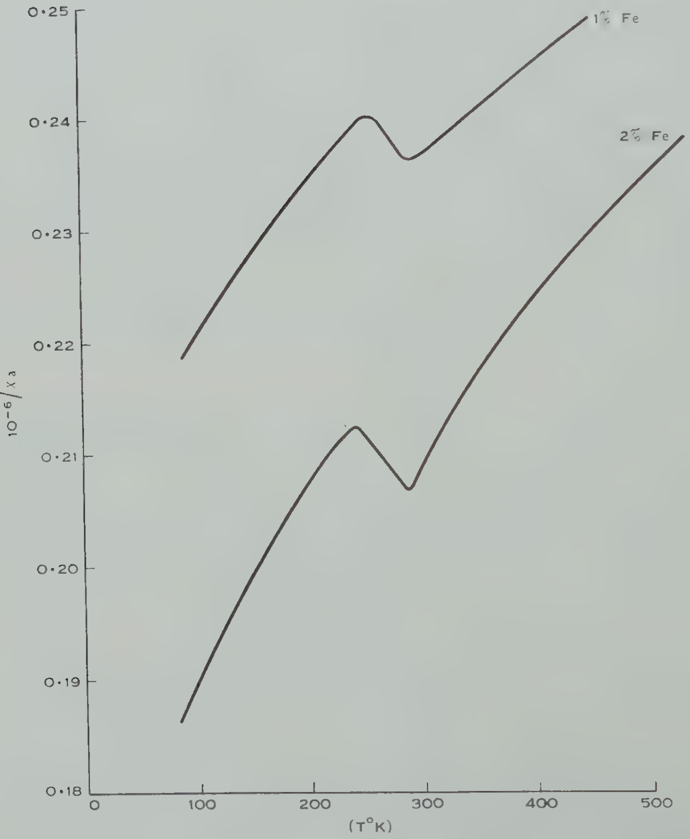


Fig. 4.—Reciprocal mass susceptibility of dilute iron alloys according to Newmann and Stevens.

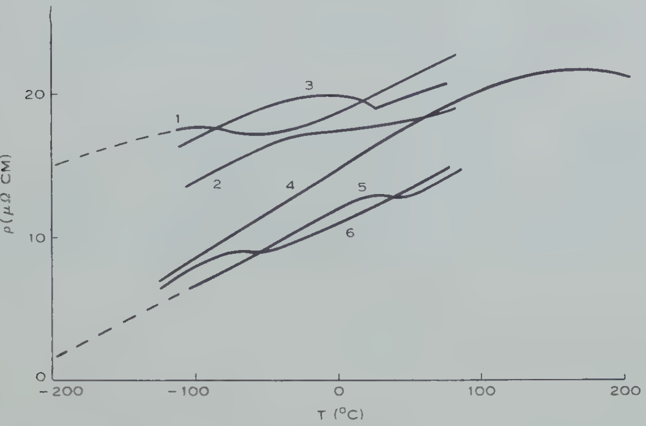


Fig. 5.—Resistivity of one per cent. alloys of various metals in chromium (after de Vries 1959). 1, Ni; 2, Co; 3, Fe; 4, Mn; 5, Cr; 6, V.

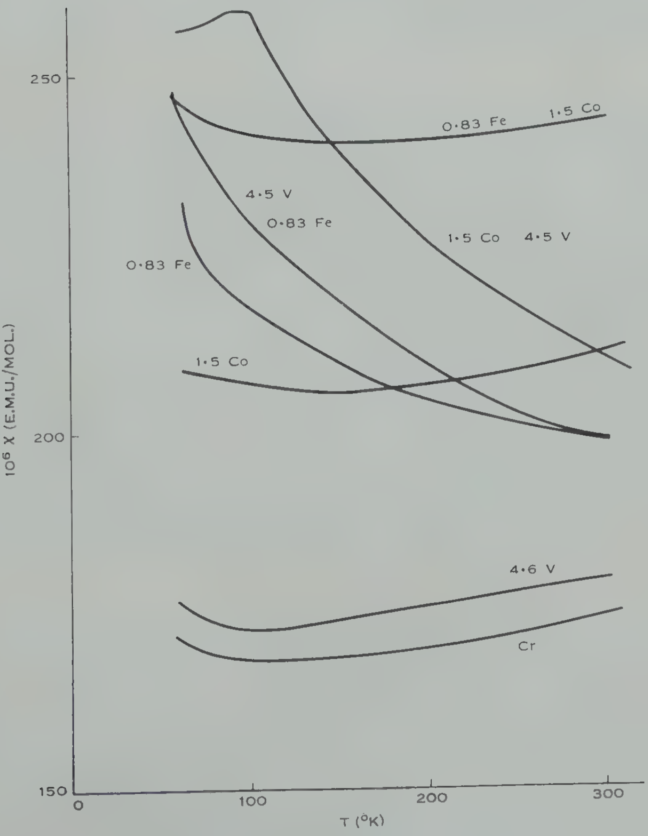


Fig. 6.—Susceptibility of some binary and ternary alloys, redrawn from the work of Lingelbach (1957).

The results for some of the ternary alloys are suggestive. The addition of 4·5 at. % V to an alloy with 1·5% Co in it converts it from the temperature-independent type to the temperature-dependent type in the range 100–300 °K. At 100 °K the measurement shows a distinct maximum and at 80 °K the susceptibility has levelled off. This behaviour would be consistent with the suggestion that Δ for cobalt in antiferromagnetic chromium is about $k \times (300\text{–}400\ ^\circ\text{K})$, and that the addition of 4·5% V decreases the chromium spins and weakens the coupling Δ to the value of about 150 °K. The results of de Vries (1959) on the variation of the resistivity anomaly temperature on alloying indeed show a very sharp drop of Neel temperature of about 90 °K for 1% V, which suggests a considerable decrease in the magnitude of the spins in chromium. de Vries's results do not allow any guess as to the Neel point of 1·5% Co 4·5% V alloy.

Another critical ternary alloy is that of composition 1·5% Co 0·83% Fe. The binary alloy with iron shows temperature dependence, the cobalt one does not, and the ternary one does not. Stevens and Newmann's work suggests the binary alloy with iron is in fact antiferromagnetic but that Δ is perhaps only 70 °K or so. The addition of 1·5% Co apparently increases the strength of this coupling up to a value of about 300 °K, so that the iron spins are effectively locked over most of the temperature range.

IV. Conclusions

The majority of the magnetic measurements reported recently on chromium-based alloys can be rationalized on the following set of assumptions.

(i) Chromium-rich alloys are usually antiferromagnetic at low temperatures. The Neel point is variable but is usually near 300 °K, except for vanadium alloys, which may have much lower Neel points.

(ii) Above the Neel point, chromium shows only a Pauli paramagnetism and magnetic atoms in solution behave in a Curie-Weiss manner.

(iii) Below the Neel point, magnetic atoms in solution have distinct energy levels according to the orientation of their spins relative to the neighbouring Cr spins. The magnitude of this splitting is large for Co (>300 °K) and small for Fe (<80 °K). At temperatures small compared with the splitting we get a temperature-independent magnetism, and at high temperatures something roughly approximating Curie-Weiss.

(iv) Interactions between different solutes are not negligible, and the treatment of Section II of this paper is very preliminary in nature.

Rather than attempt a more detailed analysis of the existing experiments, it seems clear that we should await more experiments designed specifically to test some of these interpretations. Neutron diffraction may be decisive, of course, but other magnetic experiments, e.g. extension of Stevens and Newmann's work to ternary alloys with vanadium; work with nickel (which de Vries suggests depresses the Neel point), and possibly resonance experiments to determine the splitting directly, are very desirable.

The existence of a set of narrowly separated energy levels should give a detectable specific heat anomaly of Schottky type in the region of $T=\Delta/k$. If our interpretation of the ternary 1·5% Co 4·5% V alloy behaviour is correct this should be in the easily accessible range around 100 °K.

V. Acknowledgments

I am grateful to Dr. Stevens and Dr. Newmann for allowing me to use their unpublished material, and to Dr. Bacon and Dr. Weiss for discussion on the magnetic structure of chromium metal. Many of my colleagues, especially Mr. Le Claire, Dr. Marshall, and Dr. Gardner, have been helpful in discussions.

VI. References

Lidiard, A. B. (1954).—*Rep. Prog. Phys.* **17** : 201.
Lingelbach, R. (1958).—*Z. Phys. Chem.* **14**, Pt. I and II.
Newmann, M. M., and Stevens, K. W. H. (1959).—*Proc. Phys. Soc. Lond.* **74** : 290.
Shull, C. E., and Wilkinson, M. K. (1953).—*Rev. Mod. Phys.* **25** : 100.
Sully, A. H. (1954).—" Chromium." (Butterworths Scientific Publications : London.)
de Vries, G. (1959).—*J. Phys. Radium* **20** : 438.
Wagenknecht, F. (1956).—*Z. Phys. Chem.* **5** : 332, 340.

AUSTRALASIAN MEDICAL PUBLISHING CO. LTD.
SEAMER AND ARUNDEL STS., GLEBE, SYDNEY

NOTICE TO CONTRIBUTORS

1. General.—Papers will be considered for publication irrespective of the organization to which the authors are attached. Intending contributors, in preparing papers for submission, should follow the general style adopted in this Journal. A "Guide to Authors" may be obtained on application to the Editor.

   Papers, which should be written as concisely as possible, should be double-space typed on one side of the paper only and with liberal side margins. Every page, including those carrying tables and illustrations, should be numbered. Lengthy manuscripts should be accompanied by a table of contents with headings appropriately numbered. The organization to which the author is attached should be indicated in a footnote.

   The original typescript and one carbon copy should be forwarded.

2. Summary.—A short summary, which may be paragraphed, should precede the introduction to the paper. It should be of an informative character and written in such a way that it can be used by abstracting journals without amendment.

3. References.—References are to be cited in the text by the year of publication, e.g. Thomson (1948), and are to be arranged alphabetically, giving the author's name and initials followed by the year of publication, title of the paper (if desired), title of the periodical, volume, and pages, thus:

   Thomson, G. P. (1948).—*Proc. Phys. Soc. Lond.* **61**: 403–16.

   Abbreviations of titles of periodicals should conform to those used in "A World List of Scientific Periodicals".

4. Illustrations.—Line drawings should be made with indian ink on white drawing paper (preferably bristol board), tracing paper, or graph paper. Only blue or very pale grey ruled graph paper should be used as yellow, green, and red lines are difficult to screen out. Blue linen should not be used. Lines should be sufficiently thick and symbols sufficiently large to allow of reduction to a page size of 7¾ by 5 in. (or less). It is suggested that figures be drawn approximately twice the width at which they are to be printed. Lettering and numbering should only be pencilled on the drawings as the appropriate type will be substituted before the blocks are made. Original drawings should be submitted but, where possible, photographic copies or blue-prints of the originals should also be forwarded.

   Half-tone photographs, which should be included only where essential, should be on glossy paper and show as much contrast as possible.

5. Reprints.—Individual authors receive 40 copies of each paper free, two authors receive 30 each, and three or more authors receive 20 each, but additional copies may be purchased.

CONTENTS

PAGE

The Structure of Bismuth at Low Temperatures. By C. S. Barrett 209

Thermoelectric Power of Dilute Copper and Silver Alloys. By F. J. Blatt, M. Garber, R. H. Kropschot, and B. Scott 223

Band Structure of Monovalent Metals and their Alloys. By P. G. Klemens 238

The Lattice Thermal Conductivity of Alloys. By W. R. G. Kemp and P. G. Klemens 247

Thermal Transport in Dilute Alloys. By G. K. White 255

The Intensity of Dielectric Absorption in Alkali Halides as a Function of the Concentration of Divalent Cation Impurities. By J. S. Cook and J. S. Dryden 260

High Temperature Dielectric Breakdown of Alkali Halides. By J. J. O'Dwyer 270

The Interaction of Dislocations with Boundaries and Surface Films. By A. K. Head 278

The Influence of Temperature and Strain Rate on the Flow Stress of Magnesium Single Crystals. By Z. S. Basinski 284

The Influence of Temperature and Strain Rate on the Flow Stress of Annealed and Decarburized Iron at Subatmospheric Temperatures. By Z. S. Basinski and J. W. Christian 299

The Peierls-Nabarro Force and the Deformation of Iron and Steel. By N. Louat 309

The Orientation Dependence of Work-hardening in Crystals of Face-centred Cubic Metals. By L. M. Clarebrough and M. E. Hargreaves .. 316

The Plastic Resistance of Crystals. By J. J. Gilman 327

Annealing Kinetics of Neutron-irradiated Aluminium and Copper. By T. H. Blewitt, R. R. Coltman, and C. E. Klabunde 347

The Instability of Plastic Flow of Metals at Very Low Temperatures: II. By Z. S. Basinski 354

A Low Temperature Yield Instability in Iron. By J. B. Lean .. 359

Small Angle Intensities from Double Bragg Reflections in Cold-worked Metals. By B. E. Warren 376

X-ray Measurement of Twin Faulting in Face-centred Cubic Metals. By B. E. Warren 384

Field Emission through Dielectric Layers. By R. Gomer 391

Bombardment of Metals by Inert Gas Ions. By G. J. Ogilvie 402

Nucleation and Growth of Ice Crystals upon Crystalline Substrates. By N. H. Fletcher 408

The Photographic Process as a Diatomic Reaction. By C. Candler .. 419

Zone Theory of Liquids. By R. E. B. Makinson and A. P. Roberts .. 437

Alternative Superconducting Ground States. By J. C. Fisher 446

Susceptibility of Chromium-based Alloys of Transition Elements. By W. M. Lomer 451

AUSTRALASIAN MEDICAL PUBLISHING CO. LTD.
SEAMER AND ARUNDEL STS., GLEBE, SYDNEY